

(19)日本国特許庁 (JP)

(12) 公開特許公報 (A)

(11)特許出願公開番号

特開平8-251580

(43)公開日 平成8年(1996)9月27日

(51)Int.Cl.<sup>6</sup>  
H 0 4 N 7/24

識別記号

府内整理番号

F I

H 0 4 N 7/13

技術表示箇所

Z

審査請求 未請求 請求項の数9 FD (全17頁)

(21)出願番号 特願平8-24899

(22)出願日 平成8年(1996)1月19日

(31)優先権主張番号 386992

(32)優先日 1995年1月19日

(33)優先権主張国 米国(US)

(71)出願人 390035493

エイ・ティ・アンド・ティ・コーポレーション  
AT&T CORP.

アメリカ合衆国 10013-2412 ニューヨーク  
ニューヨーク アヴェニュー オブ  
ジ アメリカズ 32

(72)発明者 ブレイン ジェフリー ハスケル  
アメリカ合衆国, 07724 ニュージャージー,  
ティントン フォールズ、グレンウッド  
ドライブ 82

(74)代理人 弁理士 三俣 弘文

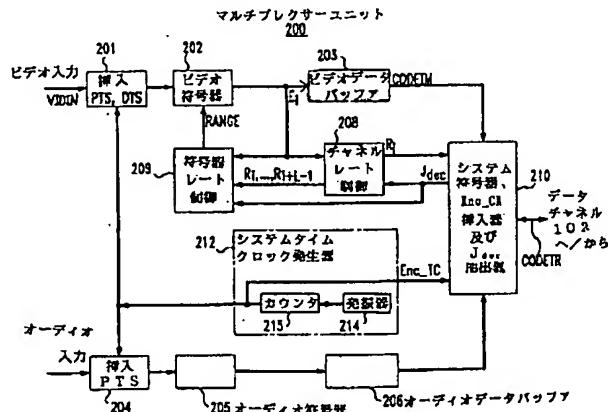
最終頁に統ぐ

(54)【発明の名称】ビデオ信号符号化装置

(57)【要約】

【課題】実時間バイトデリバリスケジュールを実復号器に対して強いるような高性能なビデオ符号器/復号器を提供する。

【解決手段】本発明のビデオ信号符号化装置は、符号化ビデオイメージ及び対応するオーディオ信号を交信するための実際に可変又は実効的に可変なビットレートのチャネルを用いるときに遭遇する符号器/復号器バッファオーバーフロー及びアンダーフローを、離れた復号器において決められたセル遅れ変化の表現、即ちジッタに応答して、ビデオ符号器のパラメーターを調節することにより克服する。



## 【特許請求の範囲】

【請求項1】 符号化ユニット(100)内に用いられ、少なくとも一つのイメージ表示を含むフレームを含むものビデオ信号を符号化し、前記符号化されたフレームを含むものビデオ信号の符号化バージョンを伝送用出力として、可変ビットレートのチャネル(110)に供給し、前記可変ビットチャネルは、前記もとのビデオ信号の符号化バージョンの蓄積用に、少なくとも一つのデコーダバッファ(302)と、もとのビデオ信号を再構成バージョンを生成するための少なくとも一つのビデオデコーダ(304)とを含むリモートデコーダユニット(102)に供給する装置において、  
 (A) 範囲表示に応答して、前記もとのビデオ信号の各フレームが、複数のビットにより表示されるような第1符号化バージョンに、前記もとのビデオ信号の各フレームを符号化する調整可能なエンコーダ(202)と、各フレーム用の前記ビットの実数は、前記範囲内に入り、前記第1の符号化フレームの各々は、前記もとのビデオ信号の第1の符号化バージョンとして、出力され、  
 (B) 前記もとのビデオ信号の第1の符号化バージョンをストアし、前記もとのビデオ信号の第1の符号化バージョンの各フレームを出力として供給するエンコーダバッファ(203)と、  
 (C) 前記遠隔デコーダユニット(102)におけるジッターの表示( $J_{dec}$ )と、前記もとのビデオ信号の前記第1符号化バージョンの少なくとも一つのフレーム内にあるビットの実数( $E_i$ )に応答し、前記チャネルに供給される現行フレーム用にビットレートリクエスト( $R_i$ )を生成し、前記もとのビデオ信号の前記符号化バージョンの所定数の将来のフレーム用に、前記チャネルにより供給されると予想される一組の予測伝送レート( $R_1, \dots, R_{M+1}$ )を生成する装置(208)と、

(D) 前記ジッター表示と、前記もとのビデオ信号の第1符号化バージョンの少なくとも一つのフレーム内のビットの実数と、前記一組の予測伝送レートに応答して、前記エンコーダバッファ手段と、前記少なくとも一つのデコーダバッファのオーバーフロー、あるいは、アンダーフローを回避するように、前記範囲を生成する装置(209)とを有することを特徴とする符号化装置。

【請求項2】 (E) エンコーダタイムクロック( $E_{nc\_TC}$ )を生成する装置(202)と、前記エンコーダタイムクロックに応答して、遠隔デコーダユニットに伝送するために、符号化信号(CODETR)内にエンコーダクロック基準信号( $E_{nc\_CR}$ )を生成する装置(210)とをさらに有することを特徴とする請求項1の装置。

【請求項3】 前記デコーダユニット(102)内に、

前記符号化信号のバージョン(CODETR')内で受信される符号化クロック基準信号を復号化する装置(301)をさらに有することを特徴とする請求項2の装置。

【請求項4】 前記デコーダユニットは、デコーダタイムクロック(Dec\_Tc)を生成する装置(308、310)をさらに有することを特徴とする請求項3の装置。

【請求項5】 前記デコーダユニットは、復号化されたエンコーダクロック基準信号と、デコーダタイムクロックの間の差分信号( $E_{nc\_CR}-Dec\_TC$ )の表示を生成する装置(310)をさらに有することを特徴とする請求項4の装置。

【請求項6】 前記デコーダユニットは、生成された差分信号に応答して、ジッター表示( $J_{dec}$ )を生成する装置(317)を有することを特徴とする請求項5の装置。

【請求項7】 前記ジッター表示は、デコーダユニット102の非同期転送モード(ATM)のセルの遅延変動を表示することを特徴とする請求項6の装置。

【請求項8】 前記デコーダユニットは、ジッター表示をエンコーダユニット(101)に送信するために、信号内に挿入する装置(301)をさらに有することを特徴とする請求項7の装置。

【請求項9】 前記エンコーダユニット(101)は、受信した信号からジッター表示を抽出する手段(210)を有することを特徴とする請求項8の装置。

## 【発明の詳細な説明】

## 【0001】

【発明の属する技術分野】 本発明は、ビデオ画像処理に関し、特に、イメージを可変ビットレートチャネル上に転送する際に符号器及び復号器のバッファオーバーフロー及びアンダーフローを避けることに関する。

## 【0002】

【従来の技術】 ISOのMPEG標準化グループは、輸送ビットストリームの実時間バイトデリバリスケジュールを実復号器に対する強いる、ビデオ符号器/復号器に対する実時間インタフェース(RTI:Real Time Interface)仕様を決めようとしている。実時間復号器モデルは、輸送バッファ、多重化バッファ及び基本ビットストリームバッファの各基本ビデオビットストリームに対する3つのバッファを含む。基本ビットストリームの各バイトがいくらかのジッタを伴って復号器に到来するであろう事実を考慮に入れてこれらのバッファの大きさは決められる。ビデオビットストリームが非同期転送モード(ATM)網上に転送される場合は、ジッタはセル遅延変化(CDV:Cell Delay Variation)として言及される。

【0003】 変換符号化技術や他の映像動画を圧縮するシステムはよく知られている。具体的には、我々による

論文、「ATM網に対する可変ビットレートビデオの制約(Constraints on Variable Bit-Rate Video for ATM Networks)」(IEEE Transactions on Circuits and Systems for Video Technology, ボリューム2、ナンバー4、361-372ページ、1992年12月)に示してある。我々はここで遅延するジッタがないと仮定して、符号器は復号器バッファ占有度の実値を計算することにより復号器バッファオーバーフロー及びアンダーフローを防げることを示した。(また1992年10月27日に我々に与えられた米国特許第5,159,447号を参照するとよい。)

#### 【0004】

【発明が解決しようとする課題】さらに、我々の標準化への寄与として、「ATM網上の可変ビットレートビデオのためのタイミングリカバリ(Timing Recovery for Variable Bit-Rate video on ATM Networks)」(MPEG92/396、AVC-315、フランス国パリ、1992年7月3日)で、付加遅延する及び付加復号器バッファ容量を取り入れることにより、ジッタの存在の下で、復号器バッファオーバーフロー及びアンダーフローを防ぐ復号器を開発した。(また1994年2月15日に我々に与えられた米国特許第5,287,182号を参照するとよい。)保証上下限をジッタに与えられて、復号器バッファのオーバーフロー及びアンダーフローは、十分に大きいバッファ及び十分に大きいジッタ遅延するを選択することにより排除できた。この解決法は、いわゆるビデオバッファ検査機構(VBV:Video Buffer Verifier)をシフトするために復号器が利用可能なバッファ容量を使って、バッファオーバーフロー又はアンダーフローがないことを保証すると考えることができる。実際は、復号器バッファはVBV揺らぎ及びジッタの両方に対応するのに十分な容量を有しない場合があり、さらに、復号器バッファは付加復号する遅延するを招き、基本ビットストリーム復号器バッファに対してだけ実用的である。本発明の目的は、実時間バイトデリバリスケジュールを実復号器に対して強いるような高性能なビデオ符号器/復号器を提供することである。

#### 【0005】

【課題を解決するための手段】符号化ビデオイメージを伝えるために実際に可変又は実効的に可変な(actually variable or effectively variable)ビットレートチャネルを用いる時に遭遇する符号器/復号器バッファオーバーフロー及びアンダーフローの問題は、復号器でセル遅延する変化(CDV)に応じて情報を生成することにより克服できる。この復号器はCDVを適応するために用いられる付加できる復号器バッファ容量を含む。この情報は符号器に転送され、ここで随意にパラメタを調整して符号化器/復号器バッファのオーバーフロー及びアンダーフローを避けられる。

#### 【0006】

この解決法の利点として、可能な最大のジ

ッタのもとでオーバーフロー及びアンダーフローを防ぐのに十分なバッファ容量で復号器を設計する必要性を軽減させることや、また復号遅延を最小にすることがある。

#### 【0007】

【発明の実施の形態】図1には本発明の、マルチプレクサーユニット100、可変ビットレートチャネル101及びデマルチプレクサーユニット102の簡略ブロック図で示す。原始ビデオ入力信号VIDINは、少なくとも1つのイメージ表現(image representation)を含むフレームを有し、マルチプレクサー100の入力として供給される。このようなビデオ信号は当分野でよく知られている。また、対応するオーディオ信号及びJdecがマルチプレクサーユニット100に供給される。後述するように、Jdecは、デマルチプレクサーユニット102内のビデオ復号器でセル遅延する変化を表し、マルチプレクサー100内のビデオ符号器により用いられて符号器/復号器バッファオーバーフロー及びアンダーフローを避けるようにパラメタを調整する。マルチプレクサー100は可変ビットレートチャネル101を通じてデマルチプレクサーユニット102に転送する出力ビットストリームCODETRとして供給する。後述するように、マルチプレクサー100からの出力として供給されるビットストリームはまた、デマルチプレクサーユニット102で、正確にビデオ及び対応するオーディオ信号を復号するために必要なオーディオ及び他の情報を含む。しかしながら、後述の議論ではCODETRはVIDINの符号化されたバージョンとして考えられる。

【0008】可変チャネル102は所望ビットレートR<sub>1</sub>で信号を送れるとする。多くの応用では、しかしながら、可変ビットレートチャネル101は、イーサーネット網や非同期転送モード(ATM)網上において所望ビットレートR<sub>1</sub>で遅れるかどうかはわからない。ATM網では、可変ビットレートチャネル101に入るビット数を制約するために使われる警察機能(policing function)又は使用パラメタ制御(UPC:usage parameter control)機能がある。本発明は、可変レートチャネル101でUPC機能を所望ビットレートR<sub>1</sub>より下に転送ビットレートを減らすことから防ぐために、チャネルレート制御デバイス208(図2)においてUPC機能をまねる。もし可変ビットレートチャネル101が所望ビットレートR<sub>1</sub>で信号を送れるかどうかの決定ができるなら、上記の我々の米国特許第5,159,447号で開示された技術を使って、可変ビットレートチャネル101の要求するをして、チャネル101から決定を受けることができる。

【0009】なんらかの実施において、可変ビットレートチャネル101それ自身は、その出力が入力に等しくなくすため、ビットの撤去を行う。このような実施においては、システム符号器、Enc\_CCR挿入器及びJ

$J_{dec}$  抽出器 210 (図2) はただ、消去のビットを記すだけであり、一部又は全てのビットの実際の消去はチャネル 101 により、もしそれが要求ビットレートが可能でないと決めるなら、行われる。代りに、可変ビットレートチャネル 101 がそれ自身でビットがビットストリーム CODEIM の形式から除外されるべきであるか決める。チャネルレート制御 208 により供給されるビットレート  $R_1$  はランダムではなく、所定の閾数の従って制約される。このようなチャネル制約機能は例えば漏れバケット機能 (leaky bucket function) のようによく知られている。特定なビットレート  $R_1$  での転送は、可変ビットレートチャネル 101 が利用できるビットレート上に制限を課すように働ける。

【0010】説明の簡潔さ及び明快さのため、単一のビデオチャネル及び対応するオーディオだけが、図2のマルチプレクサユニット 100 及び図3のデマルチプレクサユニット 102 で示してある。実際は、複数のビデオチャネル及び対応するオーディオがマルチプレクサユニット 100 での転送のために多重化するために用いられ、類似の複数のビデオチャネル及び対応するオーディオがデマルチプレクサ 102 経由で得られる。

【0011】具体的には、図2で簡略した形態で示されたマルチプレクサユニット 100 内には、ビデオチャネル挿入 PTS、DTS ユニット 201、ビデオ符号器 202 及びビデオデータバッファ 203 を有する。オーディオチャネルは挿入 PTS ユニット 204、オーディオ符号器 205 及びオーディオデータバッファ 206 を含む。ビデオデータバッファ 203 及びオーディオデータバッファ 206 からのデータは、システム符号器、Enc\_CR 挿入器及び  $J_{dec}$  抽出器 210 に供給される。システムタイムクロック発生器 212 は、発振器 214 及び出力カウンタ 215 を含む。カウンタ 215 の出力は、挿入 PTS、STS ユニット 201、PTS ユニット 204 並びにシステム符号器、Enc\_CR 挿入器及び  $J_{dec}$  抽出器 210 に供給される符号器タイムクロック  $E_{nc\_TC}$  値である。MPEG 仕様によれば、発振器 214 の周波数は  $90\text{ kHz} \pm 50\text{ ppm}$  である。さらに、発振器 214 の周波数は  $250\text{ }\mu\text{Hz}$  /秒を超えてはならない。この実施態様では、カウンタ 215 は、発振器から 214 を供給された各パルスに対して一度インクリメントされた出力として、33ビットの  $E_{nc\_TC}$  値を供給する。

【0012】ビデオ入力データは、一部又は全てのイメージ表現、即ちピクチャーフレーム上にプレゼンテーションタイムスタンプ (PTS:Presentation Time Stamp) を挿入する、挿入 PTS、DTS ユニット 201 に最初に供給される。もし PTS が非双方向予測イメージ表現 (nonB-picture:non-Bidirectionally Predicted Image Representation) 上に挿入されれば、復号タイムスタンプ (DTS:Decoding Time Stamp) もまた挿入さ

れねばならない。これらに対するイメージ表現に対して、PTS は DTS に対して符号化される双方向イメージ表現の数が 1 超える (ピクチャー周期で)。結果として得られた「タイムスタンプされた」ビデオデータは次に、それが従来の方法で符号化されるビデオ符号器 202 に渡す。順に、符号化ビデオデータはシステム符号器、Enc\_CR 挿入器及び  $J_{dec}$  抽出器 210 経由の転送を待ち受けるためにビデオデータバッファ 203 に供給される。

10 【0013】入力オーディオはいわゆるオーディオアクセスユニット (AAU:Audio Access Unit) の形態であり、PTS 値が一部又は全ての AAU 上に挿入される挿入 PTS ユニット 204 に供給される。その後、AAU は周知の方法でデジタルで符号化されるオーディオ符号器 205 に供給される。その後、オーディオデータバッファ 206 に供給され、システム符号器、Enc\_CR 挿入器及び  $J_{dec}$  抽出器 210 経由の転送を待ち受けれる。

【0014】PTS/DTS は、デマルチプレクサユニット 102 (図3) のオーディオ及びビデオデータの復号及び表示を制御するために用いられている。PTS/DTS に応答して復号及び表示は、例えば同期を維持して、デマルチプレクサユニット 102 内のデータバッファのオーバーフロー及びアンダーフローを避ける。典型的には、ユニット 201 及び 204 により挿入された PTS/DTS 値は、ある一定な値とシステムクロック発生器 212 からの瞬間  $E_{nc\_TC}$  値の和と等しい。

【0015】システム符号器、Enc\_CR 挿入器及び  $J_{dec}$  抽出器 210 は、データのパック及びパケットを形成して、データ通信チャネルに非同期に供給する。パックは複数のビデオ及びオーディオ符号器 (図示せず) からのパケットを含む。パックヘッダはまた、システムタイムクロック発生器 212 からの  $E_{nc\_TC}$  の瞬時値である符号器クロック基準 (Enc\_CR:encoder clock reference) 値を含む。MPEG 仕様によれば、パックは少なくとも 0.7 秒ごとに転送されなくてはならない。そこで、Enc\_CR 値は数秒ごとのレートで転送される。さらに、MPEG 仕様はチャネルデータレートがパックの継続時間の間は一定であると想定する。

【0016】パック内のパケットは、ビデオ又はオーディオデータを含み、もしあれば PTS/DTS 値はパケットヘッダに移す。それ故、パケット毎に複数の PTS/DTS の状況では、最初以外の全ては捨てられる。

【0017】 $R_1$  は、イメージ表現 (ピクチャー)  $i$  の符号化の間にデータチャネルに転送されるビット数を表し、このイメージ表現  $i$  は後述するように転送されるカレントイメージ表現に対してチャネルレート制御 208 により生成され、システム符号器、Enc\_CR 挿入器及び  $J_{dec}$  抽出器 210 に供給される。システム符号

器、E nc\_C R挿入器及びJ dec 抽出器210は次に、後にさらに述べる規定原理に従って実値を選択する。チャネルレート制御208及び符号器レート制御209は、デマルチプレクサー102から供給されるJ dec 及びビデオ符号器202から供給されるEに応答してビデオ符号器202のパラメータを制御するのに用いられる。

【0018】図3には、簡略ブロック図でデマルチプレクサーユニット102の詳細を示す。具体的には、復号されるビデオ情報を表示を待ち受けながら記憶されるビデオデータバッファ302に供給する、システム復号器、E nc\_C R抽出器及びJ dec 握入器301が示される。バッファ302からのビデオデータは、ビデオ表示制御器303、及びやがてはビデオ復号器304に供給される。同様に、オーディオデータは、システム復号器、E nc\_C R抽出器及びJ dec 握入器301から、表示するを待ち受けながら記憶されるオーディオデータバッファ305まで供給される。バッファ305からのオーディオデータは、オーディオ表示制御器306に及び、やがて、オーディオ復号器307に供給される。システム復号器、E nc\_C R抽出器及びJ dec 握入器301はまた、受信システムクロック基準(E nc\_C R)値を検知して、それらを許可信号とともにシステムタイムクロック発生器308に渡す。この許可信号は、受信される有効なE nc\_C R値及び開始値があることを示す。

【0019】システムタイムクロック発生器308は、位相比較器310、フィルタ311、電圧制御発振器(VCO:voltage controlled oscillator)312及び出力カウンタ314を含む位相同期ループである。システム復号器、E nc\_C R抽出器及びJ dec 握入器301からの許可信号は、有効なE nc\_C R値を受信する時に許可フィルタ311に供給される。開始値は、カウンタ314が受信されるE nc\_C R値の初期値にセットされるように供給され、初期受信パックのパケットの初期オーディオ又はビデオデータを正確に表示させる。カウンタ314からのD ec\_T C出力は、位相比較器310で受信E nc\_C R値と比較される。フィルタ311は、差E nc\_C R-D ec\_T Cを平滑化し、VCO312の周波数の制御に使われる制御電圧を生成する。

【0020】転送ジッタ遅延するの不在、及びVCO312の中心周波数がマルチプレクサーユニット100(図2)内の発振器314のものに近いと想定すると、フィルタ312の制御信号出力は小さく、ほぼ一定な値にやがて安定する。もし発振器214(図2)の周波数が上方にわずかにドリフトするなら、E nc\_C R値はわずかに増加し、VCO312に供給されている制御電圧に対して対応する増加をもたらし、これは順にその周波数を増加させる。これは、システムタイムクロック発

生器308から出力として供給されているD ec\_T C値を順にわずかに増加させ、このようにして受信E nc\_C Rの値をトラックする。同様に、発振器214の周波数の減少に対して、デマルチプレクサーユニット102でのE nc\_C R及びD ec\_T C値もまた、減少することによりお互いをトラックする。

【0021】フィルタ311により得た平滑化量、即ちE nc\_C R-D ec\_T C差動信号の平均計算の継続時間は、どれほど速くVCO312が応答できるかを決める。フィルタ311による少量の平滑化は、D ec\_T C及びE nc\_C Rの速いアラインメントを起こす。しかしながら、もしオーディオ及びビデオクロックがまた、VCO312から得られたなら、このような速いアラインメントはオーディオ及びビデオの品質に有害である。このような場合は、フィルタ311によるより多くの平滑化が必要である。

【0022】フィルタ311の利得又は同じくVCO312の入力感度はまた、カウンタ314に供給される出力パルスの周波数安定に有効である。もし利得が大きいなら、E nc\_C R及びSTCとの小さい差異がVCO312の出力で大きい周波数偏移を起こす。もし利得が大きすぎるなら、VCO312はずっと安定しない。もし利得がかなり小さいなら、たとえE nc\_C R及びD ec\_T Cがお互いに近くなくても安定動作が起こる。

【0023】オーディオ及びビデオ復号器タイミングは、上述のように、対応するオーディオ又はビデオのデータに含まれるプレゼンテーション/復号タイムスタンプ(PTS/DT S)により制御される。PTSが全てのイメージ表現上に起こると想定すると、ビデオデータバッファ302に記憶されている最も古いイメージ表現のDT S(DT SがないならPTS)は、ビデオ表示制御器203に渡される。ビデオ表示制御器303は、D ec\_T C値が最古DT Sの値まで増加するまで待つ。それは次に、ビデオデータバッファ302からの対応するイメージ表現に対する符号化ビデオデータを抽出し、それを復号するためビデオ復号器304に渡す。各ビデオイメージ表現ごとに、復号器304への入力、イメージ表現は、表示するのためにビデオ出力として供給される。ビデオ出力は、入力(例えば、Bピクチャーに対して)と同じイメージ表現であるか、又は記憶された前に復号されたピクチャー(例えば非B-ピクチャーに対して)である。

【0024】もしイメージ表現がDT Sを有しないなら、ビデオ表示制御器303は、前のイメージ表現のDT Sに仮のピクチャー継続時間をただ加えることにより推定DT S値を計算する。ビデオ復号器304は、次のイメージ表現が復号される前に、各イメージ表現の復号を常に完了する。最古の受信(又は推定)オーディオPTSがD ec\_T Cに等しければ、オーディオデータバッファ305内の次のオーディオアクセスユニットは、

オーディオ復号器307にオーディオ表示制御器306経由で供給される。復号オーディオは、オーディオ出力が表示するビデオ出力に対応して瞬時に渡される。

【0025】図6には、ビデオ表示制御器303の詳細を示し、下に述べる。

【0026】相当な転送ジッタ遅延するがあるなら、デマルチプレクサー102は正確に動作しない。これはシステムタイムクロック発生器208のフィルタ311がジッタなし状況よりずっと長い時間間隔に渡って平均しなくてはならないから結果として生じる。しかしながら、たとえこの場合でも、より重大な問題は、データ到着時がマルチプレクサー100(図2)での想定よりもかなり異なっていることである。このような事態が発生すると、デマルチプレクサー102内のビデオとオーディオの双方又は一方のデータバッファのオーバーフロー又はアンダーフローの可能性がある。

【0027】データバッファのオーバーフローするは、マルチプレクサー100内の符号器202により想定したことを越えるようにバッファのサイズを増やすことにより容易に避けられる。データバッファのアンダーフローは、復号の前にデータバッファのデータの余分に蓄積する「ジッタ遅延する」 $D_j$ の加算により軽減できる。付加記憶データはアンダーフローに対する保険の役割をする。ジッタ遅延するの上下限が保証されるなら、データバッファのアンダーフローの可能性を十分に大きいバッファ及び十分に大きい値を $D_j$ に選ぶことにより十分に排除できる。

【0028】実際の状況では、ジッタは復号器バッファ302の容量を超える。このような場合、マルチプレクサー100の符号器202(図2)は、ジッタが何であるかわからばこの作用を処理できる。この目的のため、計算ジッタユニット317(図3)は、受信E nc\_C R値及び復号器Dec\_T C値の間の相違を入力として用いてピーク $J_{max}$ の評価を計算する。位相比較器310からの差異値E nc\_C R-Dec\_T Cは計算ジッタユニット317に供給される。

【0029】図4の流れ図には、ジッタ計算の例を示す。ジッタ計算は、開始ステップ401により開始する。次にステップ402は、多分いかなる期待値よりも大きい値である仮のジッタ値に $J_{max}$ を初期化する。次に、ステップ403は比較器310からE nc\_C R-Dec\_T Cを読む。ステップ405は $J < J_{max} - \Delta$ であるかを検査し、ここで $\Delta$ はある安全因子である。もしステップ405での検査結果がNOであるなら、制御はステップ406に移される。もしステップ405の検査結果がYESであるなら( $J < J_{max} - \Delta$ )、これは $J_{max}$ が大きすぎることを表し、ステップ407において $J_{max}$ を、例えば $\Delta/2$ によって減少させ、制御はステップ406に移る。もし $J_{max} < J$ であるなら決めるステップ406の検査。もしステップ406での検査結

果がYESであるなら( $J_{max} < J$ )、これは $J_{max}$ が小さすぎることを表し、ステップ408は $J_{max}$ の値を例えればJの値に増やす。その後、ステップ409はシステム復号器、E nc\_C R抽出器及びJ dec 插入器301(図3)に供給されるJ dec の値として $J_{max}$ の値を出し、制御はステップ403に返される。同様に、システム復号器、E nc\_C R抽出器及びJ dec 插入器301はチャネル101にJ dec を供給する。

【0030】始めからの全体のプロセスは、符号器バッファ203(図2)から、符号化ビットストリームCODEIMのフレームの出力として、CODETR'の中のこのフレームに対応するバージョンまで供給し、これは復号器バッファ302(図3)により受信されてビデオ復号器303に供給され、この全体のプロセスは、可変ビットレートチャネル101のいかなる遅延するも除いてLT秒かかる。この時限LT秒において、Tはある符号化されていないビデオフレームの時限であり、Lは1以上で必ずしも整数ではないシステム遅延するパラメータであり、任意に実施者により選択される。ビットストリームCODETRの最初のビットが復号器バッファ302により受信された後に、ビデオ復号器304に最初のフレームを送る前に、復号器バッファ302によって正確にLT秒待つ。説明の明快さのため、ここに表す本発明の実施例は、Lが整数であるかのように記述するが、Lが整数ではない本発明の実施は明白である。

【0031】符号器バッファ203及び復号器バッファ302はそれぞれ所定の固定された最大サイズの $B_{max}^e$ 及び $B_{max}^d$ を有し、これらは演繹的に、ビデオ符号器202、チャネルレート制御208及び符号器レート制御209に知らされる。さらに、遅延する因子Lの値はビデオ符号器202、チャネルレート制御208、符号器レート制御209及び復号器バッファ302により演繹的に知られている。いくらかの状況ではLは、デマルチプレクサー102へとビットストリームCODETRの中で直接転送され、この遅延値の抽出はビデオ表示制御ユニット203(図2)の関連で記述する。

【0032】各イメージの符号化表現としてつくられたビット数を指定できるいかなる方法をも、ビデオ符号器202により用いられる。このような方法は当分野でよく知られている。符号器レート制御器202は、ビデオ符号器209へ信号RANGEを供給する。信号RANGEは、ビットストリームCODEIMの中に原始ビデオ信号VIDINの各フレームを符号化するときに作られる許容ビット数の範囲を表す。この例では、ビデオ符号器202が作るビット数を制約する範囲は、符号器バッファ203と復号器バッファ302のいずれもオーバーフロー又はアンダーフローしないように決定される。この範囲を決定する方法は下に記述する。ビデオ符号器202は、信号VIDINの各フレームを符号化し、これにより、フレームiを表すビットストリームCODE

IMの一部を作り、フレーム周期  $[(i-1)T, iT]$  で実際に  $E_i$  個のビットを含ませる。  $E_i$  は、フレーム  $i$  に対して信号 RANGE の中に符号器レート制御 209 により与えられた範囲の中に収まる数である。フレーム番号指標  $i$  は、所定の開始フレームに対する各フレームを示す。ビットストリーム CODEIM の一部である  $E_i$  ビットは、符号器バッファ 203 へ入力として供給され、転送される時まで、その中に記憶される。

【0033】評価チャネルビットレートのセット  $R_{i,1}, R_{i,2}, \dots, R_{i,L-1}$  のチャネルレート制御 208 による選択は、オーバーフロー又はアンダーフローする符号器バッファ 203 と復号器バッファ 302 を防ぐ必要性により制御される。この選択はまた、可変ビットレートチャネル 101 へのビットレート制約により制御される。この可変ビットレートチャネル 101 へのビットレート制約は時変する。この選択はまた、計算ジッタユニット 317 (図3) からチャネルレート制御器により受信する受信ジッタ  $J_{dec}$  の評価値により制御される。評価チャネルビットレートの選択で考えられる付加因子として各ビットを転送する費用、及びなんらかのチャネル制約 (例えば周知の漏れバケット制約) に対してはチャネルビットレートは、より厳しくないチャネルレート制約を後に作るためにより早い時間に節約できる事実がある。これは、イントラフレーム技術を使って符号化された原始ビデオ信号 VIDI N のフレームを間もなく転送する必要があるならば望ましい。

【0034】符号器レート制御 209 はチャネルレート制御 208 から、ビットレートの実値  $R_i$ 、及び後述の方法で選択する次の  $L-1$  個のフレームに対する将来のチャネルに可能性がある評価ビットレート値  $R_{i+1}, R_{i+2}, \dots, R_{i+L-1}$  を入力として受信する。符号器レート制御 209 はまた、計算ジッタユニット 217 (図2) から、受信ジッタ  $J_{dec}$  の評価する値を入力として受信する。符号器レート制御 209 はまた、各フレーム周期  $i$  に対してビデオ符号器 202 により出力として供給される実値ビット数  $E_i$  を入力として受信する。符号器レート制御 209 により出力として供給される  $E_i$  の範囲は後述の方法で決められ、符号器バッファ 203 又は復号器バッファ 302 のいずれもオーバーフロー又はアンダーフローしないようにされる。

【0035】チャネルレート制御 208 は、チャネル 101 に対するカレントチャネルビットレートの実値  $R_i$ 、及びカレントフレーム周期での出力としてビデオ符号器 202 により供給されるビット数の実値  $E_i$  を決める。評価チャネルビットレート  $R_{i+1}$  は、フレーム周期  $L$  フレームでチャネル 101 を渡って転送するために所望されるビット数を表して、それ故チャネルレート制御 208 により選択される。これらの選択チャネルビットレートのそれぞれは上述のように、符号器レート制御 209、並びにシステム符号器、ENC\_CCR 挿入器及

び  $J_{dec}$  抽出器 210 への入力として供給される。

【0036】次の理論的な説明は、チャネルレート制御 208 及び符号器レート制御 209 の詳細な動作の理解に特に可変ビットレートチャネル 101 が例に示した漏れバケット制約により制約される場合に有用である。  $E_i(t)$  は、時間  $t$  に符号器により出力したビット (又はバイト若しくはパケット) の数であるように定義される。いかなる与えられた可変ビットレートチャネル 101 の瞬間ビットレート  $R_i(t)$  は可変である。  $B^e(t)$  及び  $B^{d(e)}(t)$  は、ジッタ遅延するがないとを想定して、符号器により計算されるようにそれぞれ、符号器バッファ 203 及び復号器バッファ 302 の瞬間占有度である。符号器バッファ 203 及び復号器バッファ 302 はそれぞれ、所定の固定最大サイズ  $B_{max}^e$  及び  $B_{max}^d$  を有する。  $B_{max}^e$  を与えられ、マルチプレクサー 100 がその符号器バッファ 203 は決してオーバーフローすることはないことを保証するよう意図される。即ち、

【数1】

$$0 \leq B^e(t) \leq B_{max}^e \quad \forall t \quad (1)$$

である。復号器バッファ 302 が決してオーバーフロー又はアンダーフローすることはないことを保証するためにはビットレート  $E_i$  と  $R_i$  上に設けることを要する条件は、即ち、

【数2】

$$0 \leq B^{d(e)}(t) \leq B_{max}^d \quad \forall t \quad (2)$$

である。

【0037】この問題は間隔  $[(i-1)T, iT]$  でビット数であるように  $E_i$  ( $i = 1, 2, \dots$ ) を定義することにより離散的にされ、ここで  $T$  は、符号化されていないビデオ信号 VIDI N のある符号化されていないフレームの継続時間である。従って、

【数3】

$$E_i = \int_{(i-1)T}^{iT} E(t) dt \quad (3)$$

である。同様に、  $R_i$  は  $i$  番目のフレーム周期の間に可変ビットレートチャネル 101 上を転送するビット数、

【数4】

$$R_i = \int_{(i-1)T}^{iT} R(t) dt \quad (4)$$

であり、可変ビットレートチャネル 101 により与えられた実値ビットレートである。

【0038】この記述では、可変ビットレートチャネル 101 上への転送ビットレートの離散化は符号化ビットレート  $E_i$  のそれと全く同じであると想定されるが、他の離散化が行えることは明白である。

【0039】符号器バッファ203は、ビデオ符号器202からビットレート  $E(t)$  でビット列を受信し、ビットレート  $R(t)$  で符号化ビデオビットストリームCODEIMのビットを出力として供給する。従って、符号器バッファ203及び復号器バッファ302は、時間  $t = 0$  での開始の前に空であると想定すると、

【数5】

$$B^e(t) = \int_0^t [E(s) - R(s)] ds \quad (5)$$

であり、フレーム  $i$  の符号化後の符号器バッファ占有度は、

【数6】

$$B_i^e = B^e(iT) = \int_0^{iT} [E(s) - R(s)] ds \quad (6)$$

である。これは、

【数7】

$$B_i^e = \sum_{j=1}^i E_j - \sum_{j=1}^i R_j \quad (7)$$

として明示的に書け、

【数8】

$$B_i^e = B_{i-1}^e + E_i - R_i \quad (8)$$

として帰納的に書ける。

【0040】デマルチプレクサー102は、復号器バッファ302でデータを受信し始めた後、上述するように、復号し始める前に  $L T$  秒待つ。再び、Lは説明の明快さのために整数であると想定したが、これは必要ではない。

【0041】新しい時間指数  $\tau$  がデマルチプレクサー102に対して定められた。 $\tau$  は、復号し始める時はゼロである。よって、

【数9】

$$\tau = \tau + LT + \text{channel.delay} \quad (9)$$

であり、ここでchannel.delay（チャネル遅延）は可変ビットレートチャネル101に渡る転送の遅延時間である。マルチプレクサユニット100は、もしLが所定であるか又は明示的にパラメタとしてデマルチプレクサー102へ転送されるなら、復号器バッファ302の初期占有度  $B_i^{d(e)}$  (0) ( $\tau = 0$  である時) を計算できる。復号器バッファ302の初期占有度は、

【数10】

$$B_0^{d(e)} = \sum_{j=1}^L R_j \quad (10)$$

により与えられる。時間  $\tau = iT$  での復号器バッファ302の占有度は、

【数11】

$$B_i^{d(e)} = B_{i-1}^{d(e)} + R_{L+i} - E_i \quad (11)$$

【数12】

$$B_i^{d(e)} = B_0^{d(e)} + \sum_{j=1}^i R_{L+j} - \sum_{j=1}^i E_j \quad (12)$$

により決められる。

【0042】我々が1992年10月27日に与えられた米国特許第5,159,447号で以前に開示した配置では、符号器は、遅延するジッタを想定せずに、復号器バッファの正しい動作を保証した。しかしながら、もし遅延するジッタが存在すれば、符号器は復号器バッファの正確な占有度を知ることができない。ここでは我々は、データが早く又は遅くに復号器に到来する短期ジッタの可能性を含ませるためにバッファの動的な方程式を拡張する。最大遅延するジッタは  $J_{\max}$  に定義される。我々は復号器クロックが安定で正確であると想定し、全体を通してジッタがバッファ遅延  $L T$  よりはるかに小さいと想定する。

【0043】到来ジッタのため、符号器は正確に復号器バッファの占有度を知ることができない。符号器はしかしながら、ジッタがない場合に対応する仮のバッファ占有度を計算することができる。実値バッファ占有度はジッタのためにこれとは相違する。

【0044】 $B_i^{d(d)}$  は、時間  $i$  での実値復号器バッファ占有度であり、 $B_i^{d(e)}$  は、ジッタを想定せず、符号器により計算されるように仮のバッファ占有度であるようにし、 $R_{L+i}$  を  $i$  番目の間隔の間に復号器バッファに到来するビット数であり、 $E_i$  を  $i$  番目の間隔の間に復号器バッファを去るビット数であるようにする。

【0045】一般に、符号器は、式(11)で記述したように  $B_i^{d(e)}$  を計算する。しかしながら、もしデータが遅すぎで到来するなら、復号器バッファは、符号器が期待したよりも占有されていない、この場合、 $J_{\max} R_{L+i}$  だけ少ないデータが復号器バッファにより受信される。よって、符号器により期待したバッファ占有度  $B_i^{d(d)}$  は、実値復号器バッファ占有度  $B_i^{d(e)}$  と相違する。

【数13】

$$B_i^{d(d)} \geq B_i^{d(e)} - J_{\max} R_{L+i} \quad (13)$$

【0046】同様に、早すぎでデータ到来したなら、復号器バッファは期待されるより占有される。この場合、復号器バッファの満杯率は今は  $R_{L+i}$  である。

【数14】

$$B_i^{d(d)} \leq B_i^{d(e)} + J_{\max} R_{L+i+1} \quad (14)$$

【0047】式(13)及び式(14)は組み合わさり、符号器で計算された復号器バッファ占有度及び最大ジッタの実値を与えられると実値復号器バッファ占有度の上下限が得られる。

【0048】もし符号器が、 $J_{\max}$  の正しい値を知れるなら、

【数15】

$$J_{\max} R_{L+i} \leq B_i^{d(e)} \leq B_{\max}^d - J_{\max} R_{L+i+1} \quad (15)$$

15

を満足させることにより、復号器のバッファがアンダーフロー又はオーバーフローしないことを保証できる。

【0049】もし将来の転送ビットレート  $R_{L+i}$  が符号化時間  $i$  で未知であるなら、上限をかわりに使える。 $R_{L+i}$  がより正確にわかるとそれだけより多くの揺らぎが符号化ビットレートで許容でき、それ故、もっと良い映像品質が作られる。

【0050】実システムでは、符号器は  $J_{\max}$  の正しい値を知らないかもしれない。復号器はこの情報を計算して符号器に送らなくてはならない。復号器が符号器への最大ジッタをこれが起きるまで観測しないことはあり得る（それ故、信号を送ることが可能である）。復号器はまた、 $E_{nc\_CR}$  である全パケットが同じくジッタされるありそうもない場合において、ジッタが起こっていることを検知しない。さらに、復号器及び符号器の間になんらかの転送遅延するがある。従って、復号器バッファオーバーフロー又はアンダーフローは、符号器が  $J_{\max}$  の正しい値を受信する前に起こり得る。

【0051】1つの解決法として、復号器に大きなジッ\*

$$2J_{\max} R_{L+i} \leq B_i^{d(e)} \leq B_{\max}^d - 2J_{\max}^d R_{L+i+1} \quad (16)$$

となる。これらの差は符号器でジッタ値の過大評価を使うことにより適応させられた。

【0053】符号器バッファ203及び復号器バッファ302のアンダーフロー及びオーバーフローを防ぐのに必要な条件が一般的な可変ビットレートチャネルに対し※

$$R_i - B_{i-1}^e \leq E_i \leq B_{\max}^e + R_i - B_{i-1}^e \quad (18)$$

が必要とされる。これは、可変ビットレートチャネル101により与えられた既定のチャネルビットレート  $R_i$  に対して、ビットストリームCODEIMの各符号化フレームに供給されるビット数  $E_i$  への制約を表す。例えば、可変ビットレートチャネル101が一定なレートを有する場合、マルチプレクサー100はビデオ符号器202により用いられる符号化の品質を変えることにより符号器バッファ203をオーバーフロー又はアンダーフローから防ぐことができる。もしマルチプレクサー100が符号器バッファ203が満杯に近付いていると決めるなら、符号器バッファ203へビデオ符号器202により入力として供給されて符号化されるビットストリームCODEIMのビットレートは減らされる。このようなビットストリームCODEIMのビットレートでの減少は、ビデオ符号器202により行われる符号化の品質★

$$J_{\max} R_{i+L} \leq B_{i-1}^{d(e)} + R_{i+L} - E_i \leq B_{\max}^d - J_{\max} R_{i+L+1} \quad (19)$$

【数20】

\* タの評価する値で始めて、時間にが経るにつれ減らすことがある。さらに、もし符号器が利用可能なジッタ情報に基づいてその符号化パラメタを調整するなら、オーバーフロー又はアンダーフローの可能性は最小になる。さらに、もし復号器が正確に最大ジッタを評価するならば、オーバーフロー及び復号器問題が転送遅延間隔に対して起こり得る間に、符号器は間隔を制限することができる。

【0052】先の分析では復号器クロックが安定している正確であったと想定した。しかしながら、たとえ復号器クロックは、正しい周波数で安定しているとしても、それはおよそ  $J_{\max}$  である一定なオフセットを持った。例えば、符号器クロック基準  $E_{nc\_CR}$  を含む全てのパケットが全て  $J_{\max}$  により遅延される病的な場合には、復号器のタイムクロック  $D\_TC$  は  $J_{\max}$  遅くなる。クロックが遅いなら、復号器は符号器により期待されるより遅れてデータをそのバッファから除去し、クロックが早いなら、データは早すぎて除去される。このことが考慮に入れられるなら、式(15)は、

20 【数16】

※で示された。符号器バッファのオーバーフロー及びアンダーフローを防ぐため、式(1)及び式(8)から、

$$\begin{aligned} & \text{【数17】} \\ & 0 \leq B_{i-1}^e + E_i - R_i \leq B_{\max}^e \end{aligned} \quad (17)$$

$$\begin{aligned} & \text{【数18】} \\ & R_i - B_{i-1}^e \leq E_i \leq B_{\max}^e + R_i - B_{i-1}^e \end{aligned} \quad (18)$$

★を落とすことにより達せられる。符号化品質を落とす1つの方法として粗い量子化ステップ量を使用することがある。このような符号化品質を落とす方法は当分野でよく知られている。逆に、もし符号器バッファ203がアンダーフローすると警告するなら、ビデオ符号器202はより高いビットレートでビットストリームCODEIMを生成でき、これは、ビデオ符号器202により行われる符号化品質を増すこと、又はとして符号化シンタクスと整合性があり、ビデオ復号器304により処分できる出力詰めビットを供給することのいずれかによる。

【0054】オーバーフロー又はアンダーフローするところから復号器バッファ302を防ぐために、既定チャネルビットレートに対するビデオ符号器202のビットレートへの次の制約を課すことができる。

【数19】

$$\begin{aligned} & 17 \\ -J_{\max} R_{i+L} + R_{i+L} + B_{i-1}^{d(e)} \\ \geq E_i \geq R_{i+L} + B_{i-1}^{d(e)} - B_{\max}^d + J_{\max} R_{i+L+1} \end{aligned} \quad (20)$$

【0055】代りに、フレーム当たりのビット数は制約 \* 生む。  
されず、チャネルビットレート  $R_i$  に対して次の制約を \* 【数21】

$$\begin{aligned} E_i - B_{i-1}^{d(e)} & \leq (1 - J_{\max}) R_{i+L} \\ \leq E_i + \left[ B_{\max}^d - B_{i-1}^{d(e)} \right] & = J_{\max} R_{i+L+1} \end{aligned} \quad (21)$$

それは、  $i > L$  のときの、

$$\begin{aligned} E_{i-L} - B_{i-L-1}^{d(e)} & \leq (1 - J_{\max}) R_i \\ \leq E_{i-L} + \left[ B_{\max}^d - B_{i-L-1}^{d(e)} \right] & = J_{\max} R_{i+1} \end{aligned} \quad (22)$$

と同じであり、ここで、式(22)の左側の不等号を含む左側の不等式は、復号器バッファ302のアンダーフローが起こる条件を示し、式(22)の右側の不等号を含む右側の不等式は、復号器バッファ302のオーバーフローが起こる条件を示す。従って、Lフレーム前にビ★20

$$R_{i+1} J_{\max} \leq E_{i-L} + [B_{\max}^d - B_{i-L-1}^{d(e)}] - (1 - J_{\max}) R_i \quad (23)$$

【数24】

$$R_i \leq \left[ E_{i-L-1} + [B_{\max}^d - B_{i-L-2}^{d(e)}] - (1 - J_{\max}) R_{i-1} \right] / J_{\max} \quad (24)$$

を得るために使える。

【0057】仮想バッファの占有度を指示するカウンタ(図示せず)は、漏れバケット機能により制約される時に可変ビットレートチャネル101により維持される。フレーム周期  $i$  に対する  $R_i$  のビットは、仮想バッファ(これからは「バケット」と呼ぶ)へ入力として供給される。ビットがバケットからの出力として供給されるビットレートは、 $R_i^-$ (Rバー)ビット/フレーム周期である。バケットサイズは  $N_{\max}$  である。フレーム  $i$  が可変ビットレートチャネル101に供給された後の瞬間バ☆

$$R_i \leq N_{\max} - N_{i-1} + \bar{R} = N_{\max} - \sum_{j=1}^{i-1} R_j + i\bar{R} \quad (26)$$

である。式(26)は、可変ビットレートチャネル101に入力として供給できるビットレートへの漏れバケットの制約を定義する。

【0058】しかしながら、漏れバケット制約は唯一の制約ではない。実際は、復号器バッファ302のオーバーフローを防ぐことにより、可変ビットレートチャネル101に入力として供給できるビットレートへのより強い制約を課せる。特に、式(22)の右側に示された復

号器ビットレート制約は、式(26)の漏れバケットビ◆  $R_{i+1} \leq N_{\max} - N_{i-1} - R_i + 2\bar{R} + R_{i+1}^{UB}$  (28)

★デオ符号器202により供給された出力に対して依存するチャネルビットレート  $R_i$  への制約がある。

【0056】式(22)の右側は  $R_i$  へ付加制約、  
【数23】

☆ケット占有度は、

【数25】

$$N_i = N_{i-1} + R_i - \bar{R} = \sum_{j=1}^i R_j - i\bar{R} \quad (25)$$

30

である。高優先の項が落とされていないことを保証するため、ビットレート  $R_i$  はバケットが決してオーバーフローしないように、即ち、全ての  $i$  に対して  $N_i \leq N_{\max}$  又は、

【数26】

40◆ ットレート制約より厳しくなる。

【0059】式(22)の右側の  $R_i$  の上限は、 $R_{i+1}$  に依存するが、これは  $R_i$  を評価する前にはわからない。しかしながら、 $R_{i+1}$  の上限は、漏れバケット制約、

【数27】

$$R_{i+1} \leq N_{\max} - N_i + \bar{R} \quad (27)$$

又は、

【数28】

$$R_{i+1} \leq N_{\max} - N_{i-1} - R_i + 2\bar{R} + R_{i+1}^{UB} \quad (28)$$

から得られる。

【0060】式(22)の右側に代入すると、 $R_i$ は、\*

$$(1-J_{\max})R_i \leq E_{i-L} + B_{\max}^d - B_{i-L-1}^{d(e)} - J_{\max}(N_{\max} - N_{i-1} - R_i + 2\bar{R}) \quad (29)$$

【数30】

$$(1-2J_{\max})R_i \leq E_{i-L} + B_{\max}^d - B_{i-L-1}^{d(e)} - J_{\max}(N_{\max} - N_{i-1} + 2\bar{R}) \quad (30)$$

を満足させるように選べる。このようにして、復号器バッファオーバーフロー及びアンダーフローがジッタが  $J \approx 10$  【数31】

$$\begin{aligned} & (E_{i-L} - B_{i-L-1}^{d(e)}) / (1 - J_{\max}) \leq R_i \\ & \leq [E_{i-L} + B_{\max}^d - B_{i-L-1}^{d(e)} - J_{\max}(N_{\max} - N_{i-1} + 2\bar{R})] / (1 - 2J_{\max}) \end{aligned} \quad (31)$$

に避けられる。式(31)が実際は常にあるべきなよう  
に  $J_{\max} < 1/2$  秒と想定することは重要である。

【0061】図5には、可変ビットレートチャネル10  
1が漏れバケットビットレート制約を有しているときの  
例示のマルチプレクサー100に対するバッファ制御の  
原理が流れ図で示してある。従って、ルーチンは原始ビ  
デオ信号VIDINの最初のフレームの到来によりステ  
ップ500によって開始する。ステップ501は、フレ  
ーム  $i=1$  の符号化の前に時間  $t=0$  に変数初期化を行な  
う。 $B_i^e$  は、時間  $t=iT$  において符号器バッファ2  
03の占有度を表し、 $B_i^{d(e)}$  は、符号器によって計  
算される、復号器バッファ302の仮の完全度を表す。  
時間  $T = iT$  でジッタがないと想定して、時間  $T = iT$   
で漏れバケットの占有度を表す  $N_i$  は、全てゼロに初期  
化される。ステップ502では、原始ビデオ信号VID★

$$\begin{aligned} & (E_{j-L} - B_{j-L-1}^{d(e)}) / (1 - J_{dec}) \leq R_j \\ & \leq [E_{j-L} + B_{\max}^d - B_{j-L-1}^{d(e)} - J_{dec}(N_{\max} - N_{j-1} + 2\bar{R})] / (1 - 2J_{dec}), \end{aligned} \quad (32)$$

が得られ、ここで、式(32)の左側の不等式は、復号  
器バッファ302に対するアンダーフロー条件を示し、  
式(32)の右側の不等式は、復号器バッファ302に  
対するオーバーフロー条件を示す。

$$R_j \leq [E_{j-L-1} + B_{\max}^d - B_{j-L-1}^{d(e)} - (1 - J_{dec})R_{j-1}] / J_{dec} \quad (34)$$

【数35】

$$N_j = N_{j-1} + R_j - \bar{R} \quad (35)$$

【数36】

$$B_{j-L}^{d(e)} = B_{j-L-1}^{d(e)} + R_j - E_{j-L} \quad (36)$$

式(36)に上記の  $j$  値を代入することにより、ステッ  
プ502の仮の占有度値が得られる。

【0062】たいていの場合、 $j < i + L - 1$  に対し  
て、我々は前の評価を単純化して再利用することができ

★  $i$  の現在のフレーム  $i$  及び原始ビデオ信号VIDIN  
の次の  $L$  個の将来のフレームに対する可変ビットレート  
チャネル101から得られるビットレートの評価が決め  
られる。また、将来のビットレートの上限が決められ  
る。さらに、漏れバケット占有度及び同じ次の  $L$  個の将  
来のフレームに対する復号器バッファ302の仮の占有  
度は評価される。チャネルビットレートの不等式を評価  
するため、式(22)及び式(26)が用いられ、ここで評価  $J_{dec}$  が実値ジッタ値  $J_{\max}$  に置き換わる。 $k \leq 0$   
の場合は、 $E_k = 0$  である。漏れバケット占有度及び復  
号器バッファ302の占有度はそれぞれ、式(25)及び  
式(12)から決められる。書き直せば、 $j = i, i+1, \dots, i+L-1$  に対して、

【数32】

☆ 【数33】

$$R_j \leq N_{\max} - N_{j-1} + \bar{R} \quad (33)$$

☆ 【数34】

る、その場合評価を  $j = i + L - 1$  に対してだけする必  
要がある。しかしながら、下に示されるように、 $B_{i-L}^{d(e)}$   
は変化でき、従って、全ての評価の再評価は望ま  
しくあり得る。一般には、式(32)、式(33)及び  
式(34)のいずれかの上限と等しい  $R_j \geq 0$  の値がより  
少ないと良い選択である。しかしながら、非常に多  
くのビット数を有するフレームが起こるうとすれば、よ  
り小さい値が望ましいことが知られている。 $R_j$  のより  
小さな値はまた、チャネルが可能な限り少しありビット  
を用いないビデオディスクであったなら望ましい。 $R_j$

のより小さい値が選択されれば、この $R_i$ は復号器バッファアンダーフローを防ぐために式(32)の下限より大きいように選ばれる。

【0063】 $i \leq l$ に対しては、フレームはまだビデオ復号器304により復号されていない、復号器バッファ302はほとんど占有されていない。一般には、

$R_1, \dots, R_l$ の和は、復号器バッファ302のアン\*

$$R_{i+L} \leq UB R_{i+L} = N_{\max} - N_{i+L-1} + \bar{R} \quad (37)$$

ステップ505では、 $E_i$ の上限が符号器レート制御209により決められる。この上限は、不等式(18)から符号器バッファ203のオーバーフローへ及び不等式(19)から復号器バッファ302のアンダーフロー、の制約を用いることにより決められる。

【数38】

$$E_i \leq B_{\max}^e + R_i - B_{i-1}^e \quad (38)$$

【数39】

$$E_i \leq UBR_{i+L} + B_{i-1}^{d(e)} - J_{\max} UBR_{i+L} \quad (39)$$

これらの $E_i$ の2つの上限の最小値は、ビデオ符号器202への信号RANGEとして符号器レート制御209により出力として供給される。

【0065】ステップ506では、 $E_i$ の上下限及びビットレート $R_i$ での符号器バッファ203からのフレームの転送の支配を受けるビデオ符号器202により、フレーム*i*の符号化を始める。ステップ507では、ビデオ符号器202によるフレーム*i*の符号化、及び符号器バッファ203から供給されたフレームの可変ビットレートチャネル101を通しての転送が完了される。ビデオ符号器202はここで、チャネルビットレート制御208及び符号器レート制御209への入力として $E_i$ の実値を供給する。 $E_i$ の実値は、条件付き分岐点508で不等式(17)の評価に使われ、それにより、符号器バッファ203のアンダーフローがフレーム*i*の符号化の間に起こったかどうか決める。もしステップ508での検査結果がNOであるなら、制御はステップ510に渡される。もしステップ508での検査結果がYESであるなら、チャネルレート制御208及びチャネルインタフェース80は早く転送を終端させ、制御はステップ509に渡され、

【数40】

$$R_i = B_{i-1}^e + E_i \quad (40)$$

として $R_i$ の新しいより低い値を計算する。その後、制御はステップ510に渡される。ステップ510は、式(8)、(35)及び(36)を使ってそれぞれ $B_i^e$ 、 $N_i$ 及び $B_{i-1}^{d(e)}$ の実値を計算するために $E_i$ 及び $R_i$ の実値を使う。ステップ511では、 $i$ は次のフレームの符号化の準備でインクリメントされる。もしステップ509が実行されなかったなら、ステップ502で前に評

\* ダーフローを避けるために初期少数のビットストリームCODEIMのフレームの最初の数フレームの期待ビットレートを超えるように選ばれるべきである。

【0064】ステップ504は、式(33)で明示した漏れパケット制約を使うことにより $R_{i+L}$ の上限(UB)を評価する。

【数37】

$$R_{i+L} \leq UB R_{i+L} = N_{\max} - N_{i+L-1} + \bar{R} \quad (37)$$

10 値した全ての値はいまだ有効であり、次のフレームに対しては、ただ最近の時間間隔の評価だけが評価される必要がある。もしステップ509が実行されたなら、 $B_{i-1}^{d(e)}$ は変化し、ステップ502で前に評価した全ての値が再評価されなくてはならない。

【0066】図6には、簡略ブロック図で、ビデオ表示制御器303又はオーディオ表示制御器306のいずれかで用いられる表示制御ユニットの詳細を示す。オーディオ表示制御器306に対しては、ただPTS値だけが用いられることは重要である。具体的には、パケットヘッダを取り去り、DTSとPTSの双方又は一方の存在を検知する、パケット分解器及びDTS(PTS)抽出器601が示される。(受信された最初のパケットはPTSを含むために必要である。) PTS/DTSが検知されるなら(DTSでないならDTS=PTSをセットする)、スイッチ602は、パケット分解器及びDTS(PTS)抽出器601からの信号によりA位置にセットされ、DTS値はスイッチ602のA入力に供給される。次に、供給されたDTS値はDTS(PTS)レジスタ603へDTS1としてスイッチ602により出力として供給される。

20 【0067】パケットヘッダを取り外した後で、パケット分解器及びDTS(PTS)抽出器601がプレゼンテーションユニット(PU)検出器607への要求により、残りのパケットデータを渡す。ビデオ表示制御器303に対しては、PUはイメージ表現(即ちピクチャー)である。オーディオ表示制御器306に対しては、PUはオーディオアクセスユニット(AAU)である。全ての最初のPU(他ではない)の始めでは、プレゼンテーションユニット検出器607は、クロックDTS(PTS)レジスタ603を刻時するクロックパルスを出力する。このようにして、その入力DTS1を読まれ、記憶され、減算器604及び加算器606にDTS2である出力として供給される。次に、プレゼンテーションユニット検出器607は待つ。Dec\_TC値が減算器604に供給され、記憶されたDTS2値と比較される。結果として生じる差異Dec\_TC-DTS2はゼロ検出器605に供給される。Dec\_TCがDTS2に等しい値に増加すれば、次のPU許可信号がプレゼンテーションユニット検出器607への出力、及びセットBがスイッチ602への入力として供給される。これ

は次のPUが復号器304に渡されることを許容する。ビデオ表示制御器303に対して、この次のPU許可信号の受信において、プレゼンテーションユニット検出器304は、ビデオ復号器304へと次のPUのデータを移す。同様に、オーディオ表示制御器306に対しては、次のPU許可信号は、プレゼンテーションユニット検出器607に次のPUのデータをオーディオ復号器307に転送させる。述べたように、次のPU許可信号はまた、スイッチ602をB位置にセットさせる。もしPUの出力の間にパケット分解器及びDTS(P TS)抽出器601がPTS/D TSを検知すれば、前と同じようにスイッチ602をA位置にセットする。しかしながら、もしPTS/D TSが検知されなければ、スイッチ602はB位置に留まる。プレゼンテーションユニット検出器607によるPUデータ転送の終わりにおいて、前と同じように、これはDTS(P TS)レジスタ603のクロック入力にクロックパルスを出力する。もし新しいDTS値が到来したなら、それはスイッチ602のA入力から前と同じように読まれ、レジスタ603にDTS1値として供給されて、レジスタ603からのDTS出力値として現われる。しかしながら、もし新しいDTS値が到来しなかったなら、スイッチ602はB位置にまだ留まる。この状況では、推定DTS値は加算器606から供給され、レジスタ603への新しいDTS1値として供給され、DTS2としてその出力として現われる。この推定DTS値は、加算器606によってDec\_TC単位で一定な仮のピクチャー継続時間値を前のDTS2値に加えることにより得られる。例えば、もし仮のピクチャー周波数が29.97Hzで、Dec\_TC周波数が90Hzであれば、定数は3003である。プレゼンテーションユニット検出器607は次に、Dec\_TCがこのDTS2の新しい値に増加するまで待ち、ここで、ゼロ検出器605から次のPU許可信号(パルス)を受信する。このようにしてデータ出力サイクルを繰り返す。

## 【0068】

【発明の効果】以上述べたように、本発明のビデオ信号符号化装置は、符号化ビデオイメージ及び対応するオーディオ信号を交信するための実際に可変又は実効的に可変なビットレートのチャネルを用いるときに遭遇する符号器/復号器バッファオーバーフロー及びアンダーフローを、離れた復号器において決められたセル遅れ変化の表現、即ちジッタに応答して、ビデオ符号器のパラメーターを調節することにより克服でき、実時間バイトデリバリスケジュールを実復号器に対して強いるような高性能なビデオ符号器/復号器が提供できる。

## 【図面の簡単な説明】

【図1】本発明の可変ビットレートチャネル及びデマルチプレクサユニットを取り入れたマルチプレクサユニットの簡略ブロック図を示す。

【図2】図1のマルチプレクサユニットの詳細な簡略ブロック図を示す。

【図3】図1のデマルチプレクサユニットの詳細な簡略ブロック図を示す。

【図4】図3の計算ジッタユニットの動作を示す流れ図である。

【図5】実効的に可変ビットレートチャネルが制約されるビデオ符号器バッファ制御を示す流れ図である。

【図6】図3のビデオ表示するユニットの詳細な簡略ブロック図を示す。

## 【符号の説明】

100 マルチプレクサユニット

101 可変ビットレートチャネル

102 デマルチプレクサユニット

200 マルチプレクサユニット

201 DTSユニット

202 ビデオ符号器

203 ビデオ表示制御ユニット

204 挿入PTSユニット

205 オーディオ符号器

206 オーディオデータバッファ

208 チャネルレート制御

209 符号器レート制御

210 システム符号器、Enc\_CR挿入器及びJ<sub>dec</sub>抽出器

212 システムタイムクロック発生器

214 発振器

215 出力カウンタ

300 デマルチプレクサユニット

301 システム復号器、Enc\_CR抽出器及びJ<sub>dec</sub>挿入器

302 ビデオデータバッファ

303 ビデオ表示制御器

304 ビデオ復号器

305 オーディオデータバッファ

306 オーディオ表示制御器

307 オーディオ復号器

308 システムタイムクロック発生器

310 位相比較器

40 311 フィルタ

312 VCO

314 出力カウンタ

317 計算ジッタユニット

401 開始

402 ジッタJ<sub>max</sub>の開始値をセット

403 Enc\_CR - Dec\_TCを得る

404 絶対値J = | Enc\_CR - Dec\_TC | を計算する

405 J < J<sub>max</sub> - Δ ?

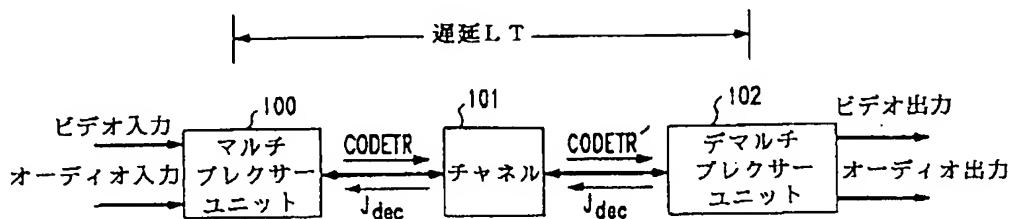
406 J<sub>max</sub> < J ?

50 407 Δ/2でJ<sub>max</sub>を減少させる

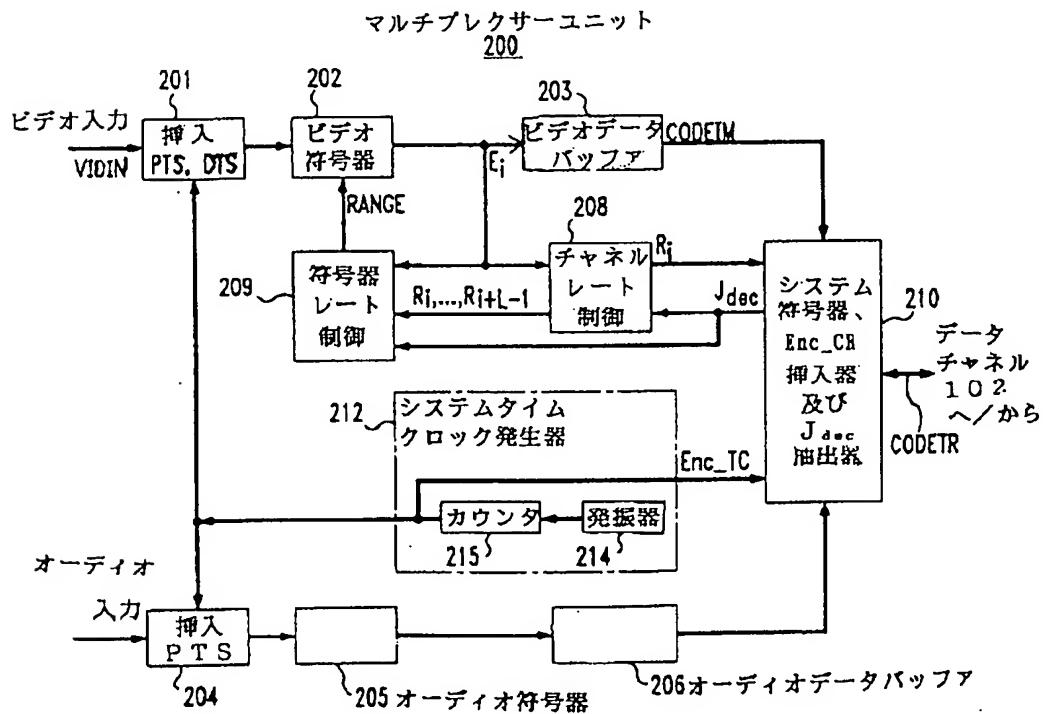
408  $J_{\max} = J$  にセットする  
 409  $J_{\max} = J_{dec}$  の値を出力  
 500 開始  
 504  $R_i + 1$  の上限を評価  
 505  $E_i$  の上限を評価  
 506 フレーム  $i$  の符号化を開始、レート  $R_i$  で転送  
 を開始  
 507 符号化及び転送を終了、 $R_i$  の実効値を得る  
 508 符号器バッファはアンダーフローしているか？\*

\* 509  $R_i$  の実効値を再計算する  
 511  $i$  をインクリメントする 次のフレームへ  
 601 パケット分解器及び DTS (PTS) 抽出器  
 602 スイッチ  
 603 DTS (PTS) レジスタ  
 604 減算器  
 605 ゼロ検出器  
 606 加算器  
 607 プレゼンテーションユニット (PU) 検出器

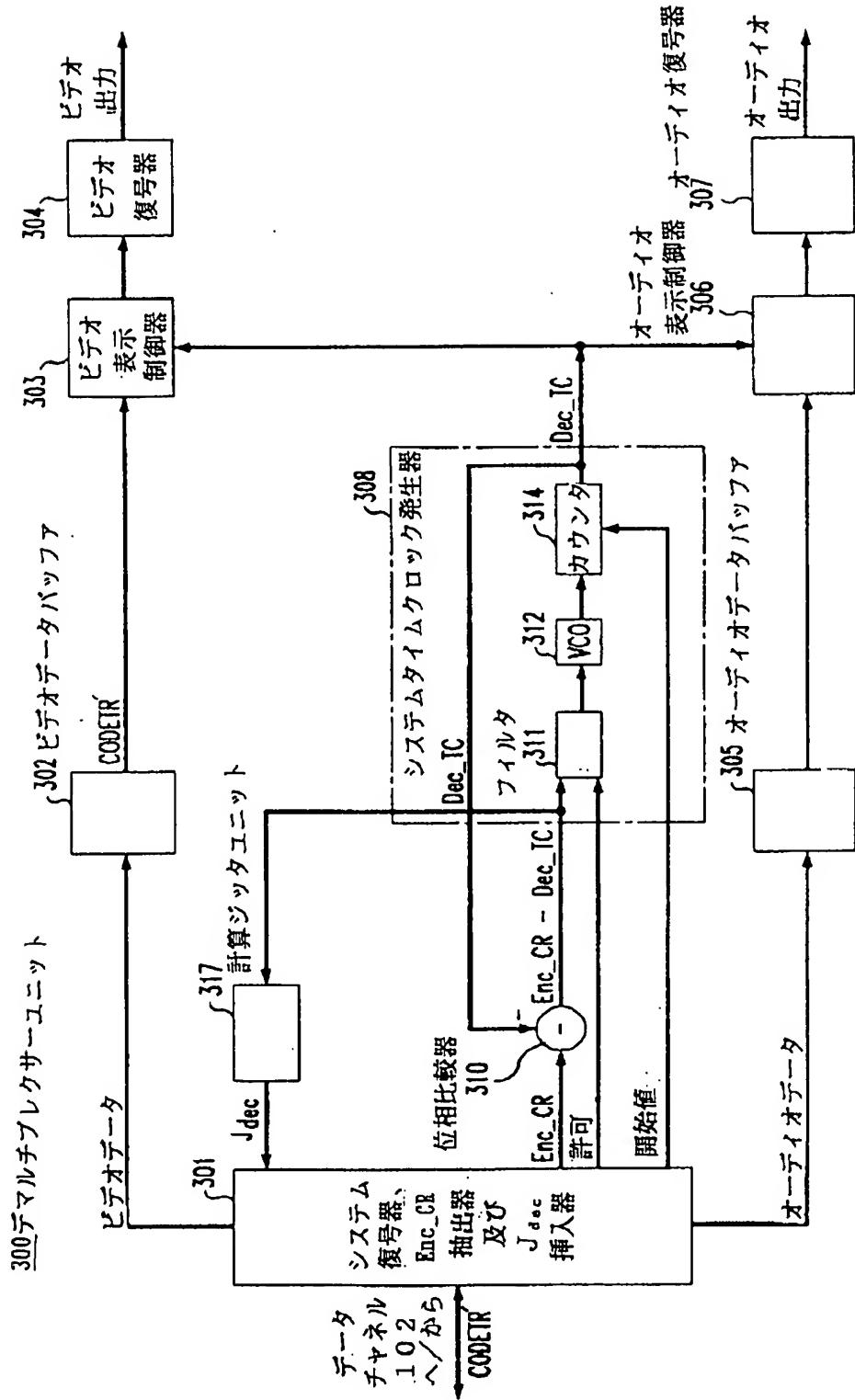
【図1】



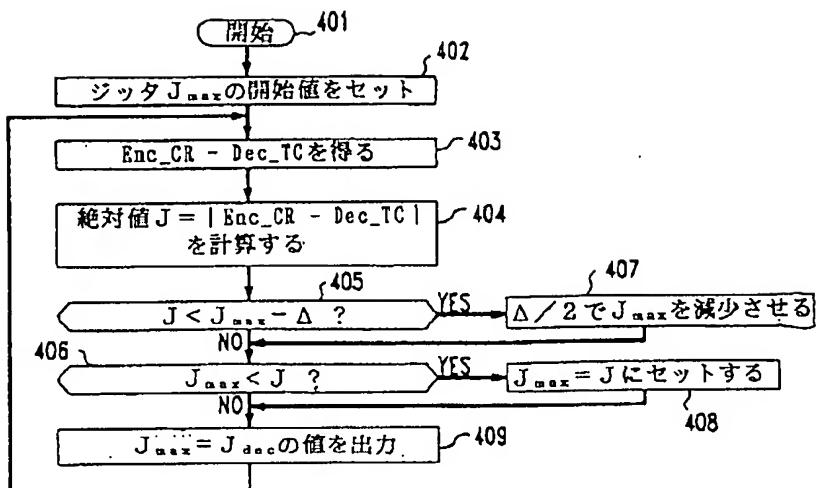
【図2】



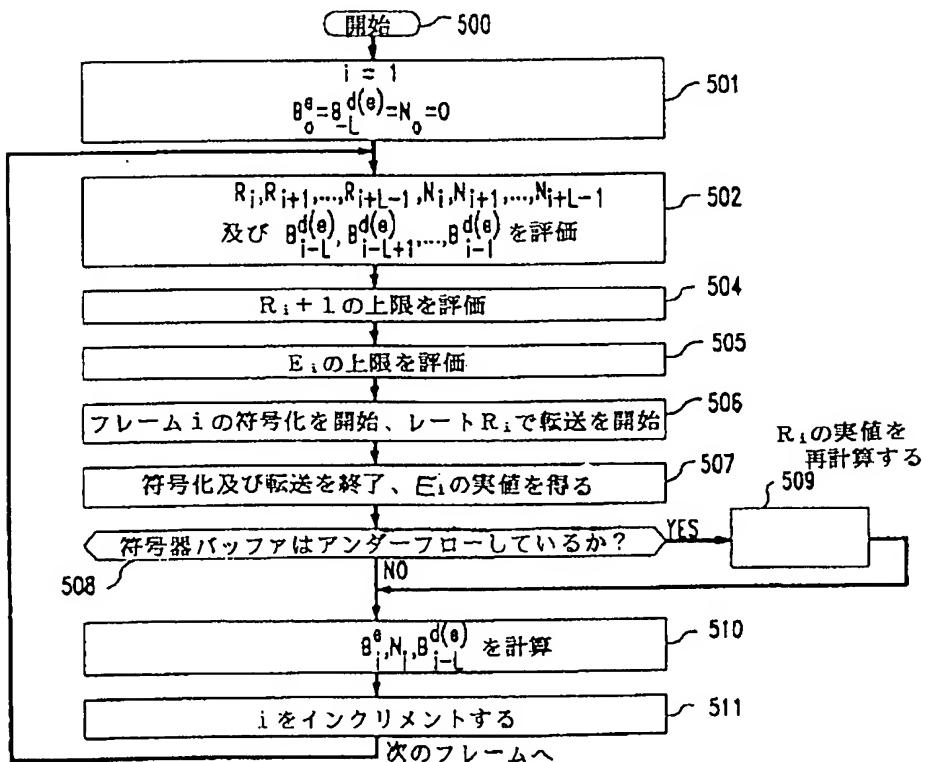
【図3】



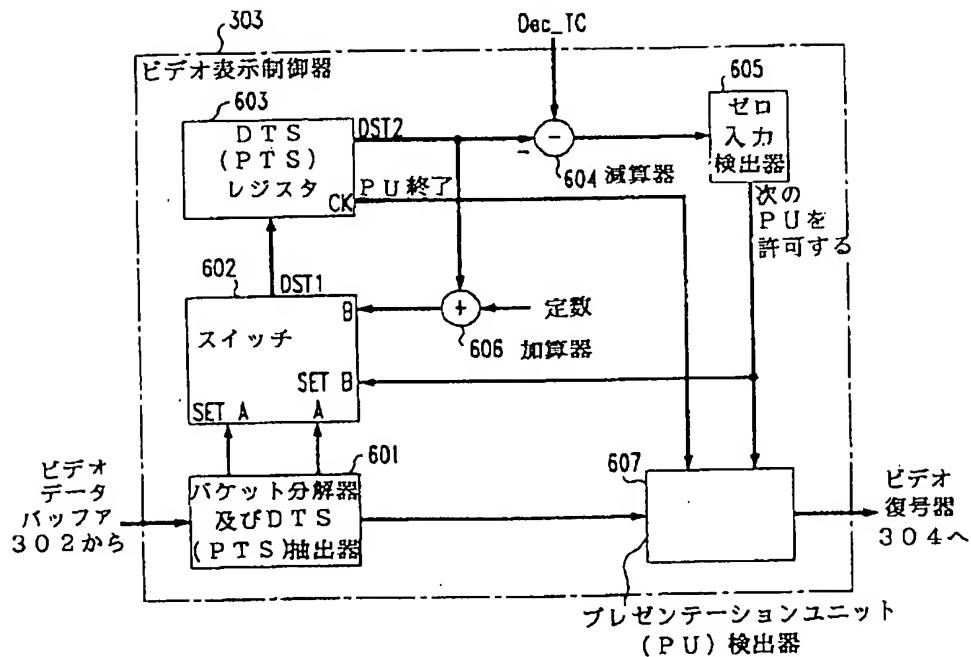
【図4】



【図5】



【図6】




---

フロントページの続き

(72)発明者 エイミー ルース レイマン  
 アメリカ合衆国, 08520 ニュージャージー,  
 イースト ウィンザー、ジェフリー  
 レーン 19

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 08-251580

(43)Date of publication of application : 27.09.1996

---

(51)Int.Cl. H04N 7/24

---

(21)Application number : 08-024899 (71)Applicant : AT & T CORP

(22)Date of filing : 19.01.1996 (72)Inventor : HASKELL BARIN G  
REIBMAN AMY R

---

(30)Priority

Priority number : 95 386992

Priority date : 19.01.1995

Priority country : US

---

(54) VIDEO SIGNAL ENCODING DEVICE

(57)Abstract:

PROBLEM TO BE SOLVED: To provide a high performance video encoder/decoder which forces a real time byte delivery schedule to a real decoder.

SOLUTION: Overflow and underflow of an encoder/decoder buffer which are encountered when a variable or effectively variable bit rate channel is actually used to communicate a coded video image and a corresponding audio signal are overcome by

responding to the representation of cell delay change that is defined by a remote decoder, i.e., jitter and adjusting the parameter of a video encoder 202.

---

[LEGAL STATUS [Date of request for examination] 08.09.1998

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number] 3168243

[Date of registration] 09.03.2001

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right] 09.03.2004

\* NOTICES \*

JPO and INPIT are not responsible for any damages caused by the use of this translation.

1. This document has been translated by computer. So the translation may not reflect the original precisely.

2.\*\*\*\* shows the word which can not be translated.

3.In the drawings, any words are not translated.

---

## CLAIMS

---

### [Claim(s)]

[Claim 1] It is used into a coding unit (100), encode the video signal of the basis containing a frame including at least one image display, and the coding version of the video signal of the basis containing said encoded frame is considered as the output for transmission. The channel (110) of a Variable Bit Rate is supplied.

Said Variable Bit Rate channel To are recording of the coding version of the video signal of said basis, at least one decoder buffer (302), In the equipment supplied to the remote decoder unit (102) containing at least one video decoder (304) for generating a reconstruction version for the video signal of a basis (A)

The encoder with which a range display is answered and each frame of the video signal of said basis encodes each frame of the video signal of said basis to the 1st coding version which is displayed by two or more bits and which can be adjusted (202), The real number of said bit for each frames goes into said within the limits. Each of said 1st coding frame It is outputted as 1st coding version of

the video signal of said basis. (B) The encoder buffer which stores the 1st coding version of the video signal of said basis, and supplies each frame of the 1st coding version of the video signal of said basis as an output (203), (C) The display of the jitter in said remote decoder unit (102) (Jdec), The real number (Ei) of the bit in at least one frame of said 1st coding version of the video signal of said basis is answered. A bit rate request (Ri) is generated on the present frames supplied to said channel. The equipment which generates the prediction transmission rate (Ri, ..., Ri+L -1) of the lot expected that the future frames of the predetermined number of said coding version of the video signal of said basis are supplied by said channel (208), (D) Said jitter display and the real number of the bit in at least one frame of the 1st coding version of the video signal of said basis, Coding equipment characterized by answering the prediction transmission rate of said lot and having said encoder buffer means and equipment (209) which generates said range so that overflow of said at least one decoder buffer or an underflow may be avoided.

[Claim 2] (E) Equipment of claim 1 characterized by having further the equipment (202) which generates an encoder timer clock (Enc\_TC), and equipment (210) which generates an encoder clock reference signal (Enc\_CR)

in a coded signal (CODETR) in order to answer said encoder timer clock and to transmit to a remote decoder unit.

[Claim 3] Equipment of claim 2 characterized by having further equipment (301) which decrypts the coding clock reference signal received within the version (CODETR') of said coded signal in said decoder unit (102).

[Claim 4] Claim 3 characterized by having further equipment (308 310) which generates a decoder timer clock (Dec\_TC) equips said decoder unit.

[Claim 5] Claim 4 characterized by having further equipment (310) which generates the display of the decrypted encoder clock reference signal and the differential signal (Enc\_CR-Dec-TC) between decoder timer clocks equips said decoder unit.

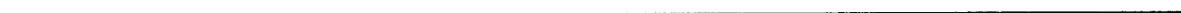
[Claim 6] Said decoder unit answers the generated differential signal, and claim 5 characterized by having equipment (317) which generates a jitter display (Jdec) equips it.

[Claim 7] Claim 6 characterized by displaying delay fluctuation of the cel of the Asynchronous Transfer Mode (ATM) of the decoder unit 102 equips said jitter display.

[Claim 8] In order to transmit a jitter display to an encoder unit (101), claim 7

characterized by having further equipment (301) inserted into a signal equips said decoder unit.

[Claim 9] Claim 8 characterized by having a means (210) to extract a jitter display from the received signal equips said encoder unit (101).



## DETAILED DESCRIPTION

---

### [Detailed Description of the Invention]

[0001]

[Field of the Invention] About a video image processing, in case especially this invention transmits an image on a Variable Bit Rate channel, it relates to avoiding the buffer overflow and the underflow of an encoder and a decoder.

[0002]

[Description of the Prior Art] The MPEG standardization group of ISO is going to decide the real-time interface (RTI:Real Time Interface) specification over the video encoder / decoder which it is forced. [ as opposed to a real decoder for the real-time cutting tool delivery schedule of a transportation bit stream ] A real-time decoder model contains three buffers to each basic video bit stream of a transportation buffer, a multiplexing buffer, and a basic bit stream buffer. These buffer sizes are decided taking into consideration the fact that each cutting tool of a basic bit stream will arrive at the decoder with some jitters. When a video bit stream is transmitted to an Asynchronous Transfer Mode (ATM) screen oversize, a jitter is mentioned as cel delay change (CDV:Cell Delay Variation).

[0003] The system which compresses a conversion coding technique and other image animations is known well. Specifically, it is shown in the paper by us, and "constraint (Constraints on Variable Bit-Rate Video for ATM Networks) of Variable Bit Rate video to an ATM network" (IEEE Transactions on Circuits and Systems for Video Technology, volume 2, 4,361 - 372 pages of numbers, December, 1992). It was shown that an encoder can prevent decoder buffer overflow and an underflow by calculating the real value of whenever [ decoder buffer occupancy ], assuming that we do not have the jitter delayed here. (It is good to refer to U.S. Pat. No. 5,159,447 given to us on October 27, 1992 again.)

[0004]

[Problem(s) to be Solved by the Invention] Furthermore, the decoder which prevents decoder buffer overflow and an underflow was indicated under the existence of a jitter as contribution to our standardization on "the timing recovery for the Variable Bit Rate video of an ATM screen oversize (Timing Recovery for Variable Bit-Rate video on ATM Networks)" MPEG 92/396, AVC-315, France country Paris, and July 3, 1992 by [ which carry out addition delay ] reaching and taking in addition decoder buffer capacity. (It is good to refer to U.S. Pat. No. 5,287,182 given to us on February 15, 1994 again.) The guarantee bound was

given by the jitter and overflow of a decoder buffer and an underflow have been eliminated by choosing a buffer large enough and jitter delay \*\*\*\* large enough. It can be considered that it guarantees that this solution does not have buffer overflow or an underflow using buffer capacity with an available decoder in order to shift the so-called video buffer verifier (VBV:Video Buffer Verifier). In practice, it may not have sufficient capacity for a decoder buffer to correspond to both VBV fluctuation and a jitter, and a decoder buffer invites delay \*\*\*\* which carries out addition decode, and is still more practical only to a basic bit stream decoder buffer. The purpose of this invention is offering highly efficient video encoder / decoder which forces it a real-time cutting tool delivery schedule to a real decoder.

[0005]

[Means for Solving the Problem] The problem of the encoder / decoder buffer overflow which encounters when actually using a strange (actually variable or effectively variable) good bit rate channel adjustable or effectually, in order to tell a coding video image, and an underflow is conquerable by generating information according to change (CDV) which carries out cel delay with a decoder. This decoder contains the addable decoder buffer capacity used since

it is adapted in CDV. This information is transmitted to an encoder, adjusts a parameter optionally here, and can avoid overflow of an encoder / decoder buffer and an underflow.

[0006] Decode delay is sometimes made [ making the need of designing a decoder with buffer capacity sufficient under the greatest possible jitter as an advantage of this solution preventing overflow and an underflow mitigate, or ] into min again.

[0007]

[Embodiment of the Invention] The simplified block diagram of the multiplexer unit 100, the Variable Bit Rate channel 101, and the demultiplexer unit 102 of this invention shows to drawing 1. The primitive video input signal VIDIN has a frame including at least one image expression (image representation), and is supplied as an input of a multiplexer 100. Such a video signal is known well for the time being in the fields. Moreover, a corresponding audio signal and corresponding Jdec are supplied to the multiplexer unit 100. Jdec expresses with the video decoder in the demultiplexer unit 102 change which carries out cel delay, and it adjusts a parameter so that it may be used by the video encoder in a multiplexer 100 and an encoder / decoder buffer overflow, and an underflow

may be avoided, so that it may mention later. A multiplexer 100 is supplied as an output bit stream CODETR transmitted to the demultiplexer unit 102 through the Variable Bit Rate channel 101. The bit stream supplied as an output from a multiplexer 100 is the demultiplexer unit 102, and includes an audio and other information required in order to decode video and a corresponding audio signal correctly again so that it may mention later. However, by the below-mentioned argument, CODETR is considered as a version by which VIDIN was encoded.

[0008] The adjustable channel 102 presupposes that a signal can be sent with the request bit rate  $R_i$ . In many application, the Variable Bit Rate channel 101 is not understood whether be behind with the request bit rate  $R_i$  in the Ethernet network or an Asynchronous Transfer Mode (ATM) screen oversize. With an ATM network, there is the police function (policing function) or use parameter control (UPC:usage parameter control) function used in order to restrain the number of bits included in the Variable Bit Rate channel 101. This invention imitates a UPC function in the channel rate control device 208 (drawing 2), in order to prevent a UPC function from reducing a transfer bit rate below the request bit rate  $R_i$  by the adjustable rate channel 101. If decision of whether the Variable Bit Rate channel 101 can send a signal with the request bit rate  $R_i$  can

be performed, using the technique indicated by our above-mentioned U.S. Pat. No. 5,159,447, demand \*\*\*\* of the Variable Bit Rate channel 101 can be carried out, and decision can be received from a channel 101.

[0009] In a certain operation, in order that the output may lose equally to an input in itself [ Variable Bit Rate channel 101 ], a bit is removed. In such operation, a system encoder, an Enc\_CR aedeagus, and the Jdec extractor 210 (drawing 2) only merely describe the bit of elimination, and actual elimination of a part or all bits will be performed, supposing it decides that a demand bit rate is not possible by the channel 101. Instead, the Variable Bit Rate channel 101 decides by itself whether the bit should be excepted from the format of a bit stream CODEIM. The bit rate  $R_i$  supplied by the channel rate control 208 is not random, and a predetermined function follows and it is restrained. Such a channel constraint function is well known like for example, a leakage bucket function (leaky bucket function). The transfer with the bit rate [ \*\*\*\* ]  $R_i$  can be committed so that a limit may be imposed on the bit rate which can use the Variable Bit Rate channel 101.

[0010] The multiplexer unit 100 of drawing 2 and the demultiplexer unit 102 of drawing 3 have shown only the single video channel and the corresponding audio for the conciseness of explanation, and lucidity. In practice, two or more

video channels and corresponding audios are used in order [ in the multiplexer unit 100 ] to multiplex for a transfer, and two or more similar video channels and corresponding audios are obtained by demultiplexer 102 course.

[0011] In the multiplexer unit 100 specifically shown with the gestalt which carried out simple by drawing 2 , it has the video channel insertion PTS, the DTS unit 201, the video encoder 202, and the video-data buffer 203. An audio channel contains the insertion PTS unit 204, the audio encoder 205, and the audio data buffer 206. The data from the video-data buffer 203 and the audio data buffer 206 are supplied to a system encoder, an Enc\_CR aedeagus, and the Jdec extractor 210. The system time clock generation machine 212 contains an oscillator 214 and the output counter 215. The output of a counter 215 is an encoder timer clock Enc\_TC value supplied to a system encoder, an Enc\_CR aedeagus, and the Jdec extractor 210 at Insertion PTS, the STS unit 201, and PTS unit 204 list. According to the MPEG specification, the frequency of an oscillator 214 is 90kHz\*\*50ppm. Furthermore, the frequency drift of an oscillator 214 must not exceed a second in 250microHz /. In this embodiment, a counter 215 supplies the Enc\_TC value of 33 bits as an output by which the increment was once carried out from the oscillator to each pulse to which 214 was supplied.

[0012] Video input data is first supplied to Insertion PTS and the DTS unit 201 which insert a presentation time stamp (PTS:Presentation Time Stamp) on a part or all image expressions, i.e., picture frames. If PTS is inserted on a non-bidirectional prediction image expression (nonB-picture:non-Bidirectionally Predicted Image Representation), a decode time stamp (DTS:Decoding Time Stamp) must also be inserted. The number of the bidirectional image expressions encoded to DTS exceeds PTS one to the image expression to these (being a picture period). Next, it passes the video data which was obtained as a result and "by which the time stamp was carried out" to the video encoder 202 encoded by the conventional approach. In order, a coding video data is supplied to the video-data buffer 203, in order to await a transfer of system encoder, Enc\_CR aedeagus, and Jdec extractor 210 course.

[0013] An input audio is the so-called gestalt of an audio access unit (AAU:Audio Access Unit), and is supplied to the insertion PTS unit 204 by which a PTS value is inserted on a part or all AAU(s). Then, AAU is supplied to the audio encoder 205 encoded in digital one by the well-known approach. Then, the audio data buffer 206 is supplied and a transfer of system encoder, Enc\_CR aedeagus, and Jdec extractor 210 course is awaited.

[0014] PTS/DTS is used in order to control decode and display of the audio of the demultiplexer unit 102 (drawing 3), and a video data. Answering PTS/DTS, decode and a display maintain a synchronization and avoid overflow and the underflow of the data buffer in the demultiplexer unit 102. Typically, the PTS/DTS value inserted by units 201 and 204 is equal to the sum of a certain fixed value and moment Enc\_TC value from the system clock generator 212.

[0015] A system encoder, an Enc\_CR aedeagus, and the Jdec extractor 210 form the pack and packet of data, and supply them to a data communication channel asynchronous. A pack contains the packet from two or more videos and audio encoders (not shown). A pack header includes the encoder clock reference (Enc\_CR:encoder clock reference) value which is the instantaneous value of Enc\_TC from the system time clock generation machine 212 again. According to the MPEG specification, a pack must be transmitted every 0.7 seconds at least. Then, an Enc\_CR value is transmitted at the rate in every several seconds. Furthermore, an MPEG specification assumes that a channel data rate is fixed between the duration of a pack.

[0016] If there is a packet in a pack including video or audio data, a PTS/DTS value will be moved to a packet header. So, in the situation of two or more

PTS/DTS, all other than the beginning are thrown away for every packet.

[0017] Ri expresses the number of bits transmitted between coding of the image expression (picture) i at a data channel, and this image expression i is generated by the channel rate control 208 to the current image expression transmitted so that it may mention later, and is supplied to a system encoder, an Enc\_CR aedeagus, and the Jdec extractor 210. A system encoder, an Enc\_CR aedeagus, and the Jdec extractor 210 choose a real value next according to the convention principle expressed further later. The channel rate control 208 and the encoder rate control 209 are used for answering Ei supplied from Jdec and the video encoder 202 which are supplied from a demultiplexer 102, and controlling the parameter of the video encoder 202.

[0018] A simplified block diagram shows the detail of the demultiplexer unit 102 to drawing 3. The system decoder, Enc\_CR extractor, and the Jdec aedeagus 301 which specifically supply the video information decoded to the video-data buffer 302 memorized while awaiting a display are shown. the video data from a buffer 302 -- the video presentation controller 303 -- and the \*\* video decoder 304 is supplied soon. Similarly, audio data are supplied to the audio data buffer 305 memorized while awaiting \*\* shown a table from a system decoder, an

Enc\_CR extractor, and the Jdec aedeagus 301. The audio data from a buffer 305 attain to the audio control indicator 306, and are soon supplied to the audio decoder 307. A system decoder, an Enc\_CR extractor, and the Jdec aedeagus 301 detect a receiving system clock criteria (Enc\_CR) value, and pass them to the system time clock generation machine 308 with an enabling signal again. It is shown that this enabling signal has the effective Enc\_CR value and effective starting value which are received.

[0019] The system time clock generation machine 308 is a phase-locked loop containing a phase comparator 310, a filter 311, a voltage controlled oscillator (VCO:voltage controlled oscillator) 312, and the output counter 314. The enabling signal from a system decoder, an Enc\_CR extractor, and the Jdec aedeagus 301 is supplied to the authorization filter 311, when receiving an effective Enc\_CR value. A starting value is supplied so that it may be set to the initial value of the Enc\_CR value with which a counter 314 is received, and it displays correctly the initial audio or video data of a packet of an initial reception pack. The Dec\_TC output from a counter 314 is compared with a receiving Enc\_CR value by the phase comparator 310. A filter 311 graduates difference Enc\_CR-Dec\_TC and generates the control voltage used for control of the

frequency of VCO312.

[0020] If it assumes that a carrying-out-transfer jitter delay absence and the center frequency of VCO312 are close to the thing of the oscillator 314 in the multiplexer unit 100 ( drawing 2 ), the control signal output of a filter 312 will be small, and will be soon stabilized in an almost fixed value. Supposing the frequency of an oscillator 214 ( drawing 2 ) carries out a drift slightly up, an Enc\_CR value increases slightly, the increment which corresponds to the control voltage currently supplied to VCO312 will be brought about, and this will make the frequency increase in order. This makes the Dec\_TC value currently supplied as an output increase from the system time clock generation machine 308 slightly in order, is carried out in this way, and carries out the truck of the value of receiving Enc\_CR. Similarly, Enc\_CR and the Dec\_TC value in the demultiplexer unit 102 also carry out the truck of each other by decreasing to reduction of the frequency of an oscillator 214.

[0021] The duration of the adjustment computation of the amount of smoothing obtained with the filter 311, i.e., an Enc\_CR-Dec\_TC differential signal, decides whether VCO312 can answer quickly however. Little smoothing with a filter 311 causes the quick alignment of Dec\_TC and Enc\_CR. However, supposing an

audio and a video clock are obtained from VCO312 again, such quick alignment is harmful to the quality of an audio and video. In such a case, many smoothing is required rather than it is based on a filter 311.

[0022] It is effective in the frequency stability of the gain of a filter 311, or the output pulse by which the input sensitivity of VCO312 is similarly supplied to a counter 314 again. Supposing gain is large, the small difference between Enc\_CR and STC will cause large frequency deviation with the output of VCO312. Supposing gain is too large, VCO312 will not be stabilized all the time. Supposing gain is quite small, even if Enc\_CR and Dec\_TC do not have near to each other, operational stability will take place.

[0023] An audio and video decoder timing are controlled by the presentation / decode time stamp (PTS/DTS) contained in the data of a corresponding audio or video as mentioned above. If it assumes that PTS happens on all image expressions, DTS (it will be PTS if there is no DTS) of the oldest image expression memorized by the video-data buffer 302 will be passed to the video presentation controller 203. It waits for the video presentation controller 303 until a Dec\_TC value increases to the value oldest [ DTS ]. It extracts the coding video data to the image expression corresponding to a degree from the

video-data buffer 302, and it passes it to the video decoder 304 in order to decode it. The input to a decoder 304 and an image expression are supplied to a displaying sake as a video outlet for every video image expression. A video outlet is the picture (as opposed to for example, an un-B-picture) decoded before being the same image expression as an input (as opposed to for example, B picture) or memorizing.

[0024] Supposing an image expression does not have DTS, the video presentation controller 303 will calculate a presumed DTS value by merely adding temporary picture duration to DTS of a pre- image expression. The video decoder 304 always completes decode of each image expression, before the following image expression is decoded. If the oldest receiving (or presumption) audio PTS is equal to Dec\_TC, the following audio access unit in the audio data buffer 305 will be supplied to the audio decoder 307 by audio control indicator 306 course. A decode audio is passed in an instant corresponding to the video outlet which an audio output displays.

[0025] The detail of the video presentation controller 303 is shown in drawing 6 , and it states below.

[0026] considerable -- although transfer jitter delay is carried out, if it is, a

demultiplexer 102 will not operate correctly. Since the filter 311 of the system time clock generation machine 208 must average over a time interval much longer than a jitter-less situation, this is produced as a result. However, a more serious problem is that data arrival time differs considerably rather than the assumption by the multiplexer 100 ( drawing 2 ) also in this case. When such a situation occurs, there is possibility of overflow of one [ the video in a demultiplexer 102 the both sides of an audio, or ] data buffer or an underflow.

[0027] It is easily avoided by increasing the size of a buffer so that what was assumed with the encoder 202 in \*\* which a data buffer overflows, and a multiplexer 100 may be exceeded. The underflow of a data buffer is mitigable with the addition of Dj which is accumulated in the excess of the data of a data buffer before decode and "which carries out jitter delay." Addition stored data carries out the role of the insurance to an underflow. If a carrying-out-jitter delay bound is guaranteed, the possibility of the underflow of a data buffer can fully be eliminated by choosing a buffer large enough and a value large enough as Dj.

[0028] In an actual situation, a jitter exceeds the capacity of the decoder buffer 302. In such a case, if the encoder 202 ( drawing 2 ) of a multiplexer 100 realizes what a jitter is, it can process this operation. The count jitter unit 317 ( drawing

3 ) calculates evaluation of Peak Jmax for this purpose, using the difference between a receiving Enc\_CR value and a decoder Dec\_TC value as an input. Difference value Enc\_CR-Dec\_TC from a phase comparator 310 is supplied to the count jitter unit 317.

[0029] The example of jitter count is shown in the flow chart of drawing 4 . Jitter count is started by the initiation step 401. Next, step 402 initializes Jmax to the temporary jitter value which is a value larger probably than what kind of expected value. Next, step 403 reads Enc\_CR-Dec\_TC from a comparator 310. Inspecting whether step 405 is  $J < J_{max} - \delta$ ,  $\delta$  is a certain insurance factor here. Supposing the inspection result in step 405 is NO, control will be moved to step 406. Supposing the inspection result of step 405 is YES ( $J < J_{max} - \delta$ ), it will express that this has too large Jmax, Jmax will be decreased by  $\delta/2$  in step 407, and it will move from control to step 406. Inspection of step 406 which will be decided supposing it is  $J_{max} < J$ . Supposing the inspection result in step 406 is YES ( $J_{max} < J$ ), it will express that this has too small Jmax, and step 408 will increase the value of Jmax to the value of J. Then, step 409 outputs the value of Jmax as a value of Jdec supplied to a system decoder, an Enc\_CR extractor, and the Jdec aedeagus 301 ( drawing 3 ), and control is returned to

step 403. Similarly, a system decoder, an Enc\_CR extractor, and the Jdec  
aedeagus 301 supply Jdec to a channel 101.

[0030] The process of the whole from the start is supplied from the encoder  
buffer 203 ( drawing 2 ) to the version corresponding to this frame in CODETR'  
as an output of the frame of the coding bit stream CODEIM, it is received by the  
decoder buffer 302 ( drawing 3 ), this is supplied to the video decoder 303, and  
the process of this whole is cut in what kind of \*\*\*\*\* LT second of the Variable  
Bit Rate channel 101 for which it is delayed. In this time limit LT second, T is the  
time limit of a certain video frame which is not encoded, and L is a parameter  
which is not necessarily an integer or more in one and which carries out system  
delay, and is chosen as arbitration by the operation person. After the bit of the  
beginning of a bit stream CODETR is received by the decoder buffer 302, before  
sending the first frame to the video decoder 304, it is LT \*\*\*\*\* correctly by the  
decoder buffer 302. Although the example of this invention expressed here is  
described for the lucidity of explanation as if L was an integer, operation of this  
invention whose L is not an integer is clear.

[0031] The encoder buffer 203 and the decoder buffer 302 have predetermined  
Bemax and predetermined Bdmax of the maximum size which were fixed,

respectively, and these are deductively told to the video encoder 202, the channel rate control 208, and the encoder rate control 209. Furthermore, the value of the delayed factor L is deductively known by the video encoder 202, the channel rate control 208, the encoder rate control 209, and the decoder buffer 302. In some situations, transfer direct of the L is carried out in a bit stream CODETR to a demultiplexer 102, and the relation of the video presentation control unit 203 (drawing 2) describes the extract of this delay value.

[0032] Any methods of specifying the number of bits built as a coding expression of each image are used with the video encoder 202. Such an approach is learned well for the time being in the fields. The encoder rate controller 202 supplies Signal RANGE to the video encoder 209. Signal RANGE expresses the range of the permission number of bits made when encoding each frame of primitive video signal VIDIN in a bit stream CODEIM. the range which restrains the number of bits which the video encoder 202 makes from this example -- both the encoder buffer 203 and the decoder buffer 302 -- overflow -- or not carrying out an underflow is determined. The approach of determining this range is described below. The video encoder 202 encodes each frame of Signal VIDIN, makes a part of bit stream CODEIM showing Frame i by this, and actually

includes the bit of E<sub>i</sub> individual with a frame period [(i-1) T, iT]. E<sub>i</sub> is a number settled into the range given by the encoder rate control 209 into Signal RANGE to Frame i. The frame number index i shows each frame to a predetermined initiation frame. E<sub>i</sub> bit which is a part of bit stream CODEIM is memorized in it until the encoder buffer 203 is supplied as an input and it is transmitted to it.

[0033] The selection by the set R<sub>i</sub> of an evaluation channel bit rate, R<sub>i+1</sub>, ..., the channel rate control 208 of R<sub>i+L</sub> -1 is controlled by the need of preventing overflow or the encoder buffer 203 which carries out an underflow, and the decoder buffer 302. This selection is controlled by bit rate constraint to the Variable Bit Rate channel 101 again. the bit rate constraint to this Variable Bit Rate channel 101 -- the time -- strange -- it carries out. This selection is controlled by the evaluation value of the receiving jitter Jdec which receives with a channel rate controller from the count jitter unit 317 ( drawing 3 ) again. In order that a channel bit rate may make behind the channel rate constraint which is not severer to the costs which transmit each bit as an addition factor considered by selection of an evaluation channel bit rate, and a certain channel constraint (for example, well-known leakage bucket constraint), there is a fact which can be saved to early time amount. this -- intra -- it is necessary to

transmit soon the frame of primitive video signal VIDIN encoded using the frame technique -- it is desirable if it becomes.

[0034] The encoder rate control 209 receives as an input evaluation bit rate value  $R_{i+1}, \dots, R_{i+L-1}$  which have possibility in the future channel to the following  $L-1$  frame chosen by the real value  $R_i$  of a bit rate, and the below-mentioned approach from the channel rate control 208. The encoder rate control 209 receives as an input the value which the receiving jitter  $J_{dec}$  evaluates from the count jitter unit 217 ( drawing 2 ) again. The encoder rate control 209 receives as an input the real value number of bits  $E_i$  supplied as an output by the video encoder 202 to each frame period  $i$  again. The range of  $E_i$  supplied by the encoder rate control 209 as an output is decided by the below-mentioned approach, an underflow is not overflowed or carried out and either the encoder buffer 203 or the decoder buffer 302 is made.

[0035] The channel rate control 208 determines the real value  $R_i$  of the current channel bit rate to a channel 101, and the real value  $E_i$  of the number of bits supplied by the video encoder 202 as an output in a current frame period. Evaluation channel bit rate  $R_{i+L}$  expresses the number of bits for which it asks in order to go across and transmit a channel 101 with  $L$  frame periods, and, so, is

chosen by the channel rate control 208. Each of these selector-channel bit rates is supplied to the encoder rate control 209 and a list as mentioned above as an input to a system encoder, an Enc\_CR aedeagus, and the Jdec extractor 210.

[0036] Especially the following theoretical explanation is useful when the Variable Bit Rate channel 101 is restrained by understanding of detailed actuation of the channel rate control 208 and the encoder rate control 209 by leakage bucket constraint shown for the example.  $E(t)$  is defined as being the number of the bits (or a cutting tool or a packet) outputted to time amount  $t$  with the encoder. Moment bit rate [ of what kind of given Variable Bit Rate channel 101 ]  $R(t)$  is adjustable. Although  $B_e(t)$  and  $B_d(e)$ , and  $(t)$  carry out jitter delay, as calculated by the encoder supposing \*\* which is not, they are whenever [ encoder buffer 203 and decoder buffer 302 moment occupancy-], respectively.

The encoder buffer 203 and the decoder buffer 302 have predetermined fixed maximum size  $B_{max}$  and predetermined  $Bd_{max}$ , respectively.  $B_{max}$  can be given, and it is meant so that it may guarantee that a multiplexer 100 never overflows the encoder buffer 203. Namely, [Equation 1]

$$0 \leq B_e(t) \leq B_{max} \quad \forall t \quad (1)$$

It comes out. They are the conditions which require preparing on a bit rate  $E_i$  and  $R_i$  in order to guarantee overflow or never not carrying out [ of 302 decoder buffer ] an underflow, i.e., [Equation 2].

$$0 \leq B^{d(e)}(t) \leq B_{\max}^d \quad \forall t \quad (2)$$

It comes out.

[0037] This problem is made discrete by defining  $E_i$  ( $i = 1, 2, \dots$ ) as being the number of bits at spacing  $[(i-1)T, iT]$ , and  $T$  is the duration of the frame with video signal VIDIN which is not encoded which is not encoded here. Therefore, [Equation 3]

$$E_i = \int_{(i-1)T}^{iT} E(t) dt \quad (3)$$

It comes out. Similarly,  $R_i$  is the number of bits and [Equation 4] which transmit the Variable Bit Rate channel 101 top between the  $i$ -th frame periods.

$$R_i = \int_{(i-1)T}^{iT} R(t) dt \quad (4)$$

It is the real value bit rate which came out, exists and was given by the Variable Bit Rate channel 101.

[0038] Although it is assumed in this description that discretization of the transfer bit rate to the Variable Bit Rate channel 101 top is completely the same as it of the coding bit rate  $E_i$ , it is clear that other discretization can be performed.

[0039] The encoder buffer 203 receives a bit string by bit rate  $E(t)$  from the video encoder 202, and supplies the bit of the coding video bit stream CODEIM as an output by bit rate  $R(t)$ . Therefore, the encoder buffer 203 and the decoder buffer 302 are [Equation 5], when it assumes that it is empty before initiation by time amount  $t=0$ .

$$B^e(t) = \int_0^t [E(s) - R(s)] ds \quad (5)$$

Come out, it is and whenever [ after coding of Frame  $i$  / encoder buffer occupancy ] is [Equation 6].

$$B^e_i = B^e(iT) = \int_0^{iT} [E(s) - R(s)] ds \quad (6)$$

It comes out. This is [Equation 7].

$$B^e_i = \sum_{j=1}^i E_j - \sum_{j=1}^i R_j \quad (7)$$

It carries out, can write clearly and is [Equation 8].

$$B_i^e = B_{i-1}^e + E_i - R_i \quad (8)$$

It carries out and can write inductively.

[0040] A demultiplexer 102 is LT \*\*\*\*\*, before beginning to decode so that it may mention above after the decoder buffer 302 begins to receive data. Again, this is not required although L assumed that it was an integer for the lucidity of explanation.

[0041] The new time amount characteristic tau was defined to the demultiplexer 102. tau is zero when beginning to decode. Therefore, [Equation 9]

$$t = \tau + LT + \text{channel.delay} \quad (9)$$

It comes out, and it is and channel.delay (channel delay) is the time delay of the transfer over the Variable Bit Rate channel 101 here. L is predetermined, or if the multiplexer unit 100 is clearly transmitted to a demultiplexer 102 as a parameter, it can calculate Bd (e), (0), and (the time of being tau= 0) whenever [ decoder buffer 302 initial occupancy-]. Whenever [ decoder buffer 302 initial occupancy-] is [Equation 10].

$$B_0^{d(e)} = \sum_{j=1}^L R_j \quad (10)$$

It is alike and is given more. Whenever [ decoder buffer 302 in time amount tau=iT occupancy-] is [Equation 11].

$$B_i^{d(e)} = B_{i-1}^{d(e)} + R_{L+i} - E_i \quad (11)$$

[Equation 12]

$$B_i^{d(e)} = B_0^{d(e)} + \sum_{j=1}^i R_{L+j} - \sum_{j=1}^i E_j \quad (12)$$

It is alike and is decided more.

[0042] In the arrangement which we indicated by U.S. Pat. No. 5,159,447 given on October 27, 1992 before, the encoder guaranteed right actuation of a decoder buffer, without assuming the delayed jitter. However, if the delayed jitter exists, an encoder cannot know whenever [ decoder buffer exact occupancy-]. here, in order for us to include the possibility of the short-term jitter which data are late [ early or ] alike and arrives at a decoder, the dynamic equation of a buffer is extended. The jitter which carries out the maximum delay is defined as Jmax. We assume that a decoder clock is stable and that it is exact, and assume that a jitter is far smaller than the buffer delay LT through the whole.

[0043] An encoder cannot know whenever [ decoder buffer occupancy-] correctly

because of an arrival jitter. Carrying out encoder measles, when there is no jitter, whenever [ temporary buffer occupancy / ] can be calculated. Whenever [ real value buffer occupancy ] is different from this for a jitter.

[0044] It is made for  $Bd(e)_i$  to be whenever [ temporary buffer occupancy ], as a jitter is not assumed but it is calculated by the encoder, it is the number of bits which arrives  $RL+i$  at a decoder buffer between the  $i$ -th spacing, and a decoder buffer is made  $Bd(d)_i$  be whenever [ by time amount  $i$  / real value decoder buffer occupancy ], and to be last number of bits between the  $i$ -th spacing about  $Ei$ .

[0045] Generally, an encoder calculates  $Bd(e)_i$ , as the formula (11) described. however, data are late -- it will elapse, it comes out, and if it comes, an encoder will expect a decoder buffer, reliance will not be occupied, either and data only with little  $J_{max}RL+i$  will be received by the decoder buffer in this case. Therefore,  $Bd(e)_i$  is [ whenever / buffer occupancy / which was expected with the encoder / whenever / real value decoder buffer occupancy ] different from  $Bd(d)_i$ .

[Equation 13]

$$B_i^{d(d)} \geq B_i^{d(e)} - J'_{max} R_{L+i} \quad (13)$$

[0046] similarly, early -- if it elapses and data arrival is come out and carried out,

a decoder buffer is occupied rather than it is expected. In this case, the rate of full of a decoder buffer is  $RL+i+1$  now.

[Equation 14]

$$B_i^{d(d)} \leq B_i^{d(e)} + J_{\max} R_{L+i+1} \quad (14)$$

[0047] A formula (13) and a formula (14) combine, and if the real value of whenever [ decoder buffer occupancy / which was calculated with the encoder ], and the maximum jitter can be given, the bound of whenever [ real value decoder buffer occupancy ] will be obtained.

[0048] It will be [Equation 15] if an encoder is found in the right value of  $J_{\max}$ .

$$J_{\max} R_{L+i} \leq B_i^{d(e)} \leq B_{\max}^d - J_{\max} R_{L+i+1} \quad (15)$$

By making it satisfied, the buffer of a decoder can guarantee an underflow or not overflowing.

[0049] By the coding time amount  $i$ , if future transfer bit rate  $RL+i+1$  is strange, it can use an upper limit instead. If accuracy understands  $RL+i+1$  more, so more much fluctuation can approve with a coding bit rate, and so better image quality will be made.

[0050] In a real system, an encoder may not know the right value of Jmax. A decoder must calculate this information and must send it to an encoder. A decoder cannot observe the maximum jitter to an encoder until this occurs (it is possible to send a signal so). all the packets that are Enc\_CR are the same and the jitter of the decoder is carried out again -- it is not likely to be -- it does not detect that the jitter has happened to the case. furthermore, between a decoder and encoders -- being certain -- although transfer delay is carried out, it is. Therefore, decoder buffer overflow or an underflow may happen, before an encoder receives the right value of Jmax.

[0051] It may begin with the value which a big jitter evaluates to a decoder as one solution, and may reduce along with \*\*\*\*\* to time amount. Furthermore, if the coding parameter is adjusted based on jitter information with an available encoder, the possibility of overflow or an underflow will become min. Furthermore, as for an encoder, spacing can be restricted while overflow and a decoder problem may arise to transfer delay spacing, supposing a decoder evaluates the maximum jitter correctly.

[0052] In previous analysis, the decoder clock was stable and it was assumed that it was exact. However, even if the decoder clock was stable on the right

frequency, it was able to have the fixed offset which is  $J_{max}$  about. the morbid case where all the packets of all containing for example, encoder clock reference Enc\_CR are delayed by  $J_{max}$  -- timer clock D\_TC of a decoder --  $J_{max}$  -- it becomes late. If a clock is late, a decoder will be behind rather than it is expected by the encoder, data are removed from the buffer, and if a clock is early, data will be too early and will be removed. A formula (15) will be [Equation 16] if put into this by consideration.

$$2J_{max} R_{L+i} \leq B_i^{d(e)} \leq B_{max} - 2J_{max}^d R_{L+i+1} \quad (16)$$

It becomes. These differences were fitted by using overestimation of a jitter value with an encoder.

[0053] Conditions required to prevent the underflow of the encoder buffer 203 and the decoder buffer 302 and overflow were shown to the general Variable Bit Rate channel. In order to prevent overflow of an encoder buffer and an underflow, it is [Equation 17] from a formula (1) and a formula (8).

$$0 \leq B_{i-1}^e + E_i - R_i \leq B_{max}^e \quad (17)$$

[Equation 18]

$$R_i - B_{i-1}^e \leq E_i \leq B_{max}^e + R_i - B_{i-1}^e \quad (18)$$

\*\* -- it is supposed that it is required. This expresses the constraint to the number of bits  $E_i$  supplied to each coding frame of a bit stream CODEIM to the fixed channel bit rate  $R_i$  given by the Variable Bit Rate channel 101. For example, when it has a rate with the fixed Variable Bit Rate channel 101, a multiplexer 100 can protect the encoder buffer 203 from overflow or an underflow by changing the quality of coding used by the video encoder 202. Supposing a multiplexer 100 decides that the encoder buffer 203 is approaching to the limit, the bit rate of the bit stream CODEIM which is supplied to the encoder buffer 203 as an input by the video encoder 202, and is encoded will be reduced. The reduction with such a bit rate of a bit stream CODEIM is attained by lowering the quality of coding performed by the video encoder 202. The amount of quantization steps coarse as one method of lowering coding quality may be used. The method of lowering such coding quality is learned well for the time being in the fields. On the contrary, if it warns that the encoder buffer 203 carries out an underflow, the video encoder 202 can generate a bit stream CODEIM with a higher bit rate, and this will be based on either of supplying increasing the coding quality performed by the video encoder 202, and the output stuffing bit which a pigeon is carried

out, and there are coding syntax and adjustment, and can be disposed of with the video decoder 304.

[0054] Overflow or since an underflow is carried out, in order to prevent the decoder buffer 302, constraint next to the bit rate of the video encoder 202 to a fixed channel bit rate can be imposed.

[Equation 19]

$$J_{\max} R_{i+L} \leq B_{i-1}^{d(e)} + R_{i+L} - E_i \leq B_{\max}^d - J_{\max} R_{i+L+1} \quad (19)$$

[Equation 20]

$$\begin{aligned} & -J_{\max} R_{i+L} + R_{i+L} + B_{i-1}^{d(e)} \\ & \geq E_i \geq R_{i+L} + B_{i-1}^{d(e)} - B_{\max}^d + J_{\max} R_{i+L+1} \end{aligned} \quad (20)$$

[0055] Instead, the number of bits per frame is not restrained, but induces the next constraint to the channel bit rate  $R_i$ .

[Equation 21]

$$\begin{aligned} E_i - B_{i-1}^{d(e)} & \leq (1 - J_{\max}) R_{i+L} \\ & \leq E_i + \left[ B_{\max}^d - B_{i-1}^{d(e)} \right] - J_{\max} R_{i+L+1} \end{aligned} \quad (21)$$

It is [Equation 22] at the time of  $i > L$ .

$$E_{i-L} - B_{i-L-1}^{d(e)} \leq (1 - J_{\max}) R_i \\ \leq E_{i-L} + [B_{\max}^d - B_{i-L-1}^{d(e)}] - J_{\max} R_{i+1} \quad (22)$$

It is the same, and the inequality of the left-hand side containing the greater than sign on the left-hand side of a formula (22) shows the conditions from which the underflow of the decoder buffer 302 arises, and the inequality of the right-hand side containing the greater than sign on the right-hand side of a formula (22) shows the conditions from which overflow of the decoder buffer 302 takes place here. Therefore, there is constraint to the channel bit rate  $R_i$  for which it depends to the output supplied by the video encoder 202 L-frame ago.

[0056] The right-hand side of a formula (22) is addition constraint and [Equation 23] to  $R_i$ .

$$R_{i+1} J_{\max} \leq E_{i-L} + [B_{\max}^d - B_{i-L-1}^{d(e)}] - (1 - J_{\max}) R_i \quad (23)$$

[Equation 24]

$$R_i \leq [E_{i-L-1} + [B_{\max}^d - B_{i-L-2}^{d(e)}] - (1 - J_{\max}) R_{i-1}] / J_{\max} \quad (24)$$

It can use for a \*\*\*\*\* sake.

[0057] The counter (not shown) which directs whenever [ virtual buffer

occupancy-] is maintained by the Variable Bit Rate channel 101, when restrained by the leakage bucket function. The bit of  $R_i$  to frame period  $i$  is supplied to a virtual buffer (from now on, it is called a "bucket") as an input. The bit rates to which a bit is supplied as an output from a bucket are  $R \sim (R \bar{R})$  bit / frame period. Bucket size is  $N_{max}$ . Whenever [ moment bucket occupancy / after Frame  $i$  was supplied to the Variable Bit Rate channel 101 ] is [Equation 25].

$$N_i = N_{i-1} + R_i - \bar{R} = \sum_{j=1}^i R_j - i\bar{R} \quad (25)$$

It comes out. They are [ as opposed to / in order to guarantee that the term of high priority is not dropped, so that a bucket may never overflow a bit rate  $R_i$  / all  $i$  ] nickel  $\leq N_{max}$  or [Equation 26].

$$R_i \leq N_{max} - N_{i-1} + \bar{R} = N_{max} - \sum_{j=1}^{i-1} R_j + i\bar{R} \quad (26)$$

It comes out. A formula (26) defines constraint of the leakage bucket to the bit rate which can be supplied as an input as the Variable Bit Rate channel 101. [0058] However, leakage bucket constraint is not the only constraint. in practice, impose the stronger constraint to the bit rate which can be supplied to the Variable Bit Rate channel 101 as an input by preventing overflow of the decoder

buffer 302 -- \*\*. Especially the decoder bit rate constraint shown in the right-hand side of a formula (22) becomes severer than leakage bucket bit rate constraint of a formula (26).

[0059] Although it depends for the upper limit of  $R_i$  on the right-hand side of a formula (22) on  $R_{i+1}$ , before this evaluates  $R_i$ , it is not understood. However, the upper limit of  $R_{i+1}$  is leakage bucket constraint and [Equation 27].

$$R_{i+1} \leq N_{\max} - N_i + \bar{R} \quad (27)$$

Or [Equation 28]

$$R_{i+1} \leq N_{\max} - N_{i-1} - R_i + 2\bar{R} + R_{i+1}^{UB} \quad (28)$$

since -- it is obtained.

[0060]  $R_i$  is [Equation 29] when it substitutes for the right-hand side of a formula (22).

$$(1 - J_{\max}) R_i \leq E_{i-L} + B_{\max}^d - B_{i-L-1}^{d(e)} - J_{\max} (N_{\max} - N_{i-1} - R_i + 2\bar{R}) \quad (29)$$

[Equation 30]

$$(1 - 2J_{\max}) R_i \leq E_{i-L} + B_{\max}^d - B_{i-L-1}^{d(e)} - J_{\max} (N_{\max} - N_{i-1} + 2\bar{R}) \quad (30)$$

It can choose so that it may be made satisfied. Thus, it is [Equation 31] when decoder buffer overflow and the underflow of a jitter are smaller than Jmax.

$$\begin{aligned} & (E_{i-L} - B_{i-L-1}^{d(e)}) / (1 - J_{\max}) \leq R_i \\ & \leq \left[ E_{i-L} + B_{\max}^d - B_{i-L-1}^{d(e)} - J_{\max} (N_{\max} - N_{i-1} + 2\bar{R}) \right] / (1 - 2J_{\max}) \end{aligned} \quad (31)$$

It is alike and is avoided. It is important to assume that there should always be a formula (31) in practice with  $J_{\max} < 1 / 2$  seconds.

[0061] The flow chart has shown the principle of the buffer control to the multiplexer 100 of instantiation when the Variable Bit Rate channel 101 has leakage bucket bit rate constraint to drawing 5. Therefore, a routine is started by step 500 by arrival of the frame of the beginning of primitive video signal VIDIN. Step 501 carries out variable initialization to time amount  $t = 0$  before coding of a frame  $i = 1$ . Bei -- time amount whenever [ encoder buffer 203 occupancy-] is expressed in  $t = iT$  --  $B_d(e) - L$  expresses whenever [ of the decoder buffer 302 calculated by the encoder / temporary / full ]. All nickel that assumes that there is no jitter at time amount  $T = iT$ , and expresses whenever [ leakage bucket occupancy-] with time amount  $T = iT$  is initialized by zero. It opts for evaluation of the bit rate obtained from the Variable Bit Rate channel 101 to the current frame i

of primitive video signal VIDIN, and L future frames next to primitive video signal VIDIN at step 502. Moreover, the upper limit of a future bit rate is decided. Furthermore, whenever [ leakage bucket occupancy ], and whenever [ decoder buffer 302 to following L same future frames temporary occupancy-] are evaluated. In order to evaluate the inequality of a channel bit rate, a formula (22) and a formula (26) are used and Evaluation Jdec replaces the real value jitter value Jmax here. In the case of k<=0, it is Ek=0. Whenever [ leakage bucket occupancy ], and whenever [ decoder buffer 302 occupancy-] are decided from a formula (25) and a formula (12), respectively. If it rewrites, it will be [Equation 32] to j=i, i+1, ..., i+L-1.

$$\begin{aligned} & (E_{j-L} - B_{j-L-1}^{d(e)}) / (1 - J_{dec}) \leq R_j \\ & \leq [E_{j-L} + B_{max}^d - B_{j-L-1}^{d(e)} - J_{dec}(N_{max} - N_{j-1} + 2\bar{R})] / (1 - 2J_{dec}), \end{aligned} \quad (32)$$

The inequality on the left-hand side of a formula (32) shows the underflow conditions over the decoder buffer 302, and the inequality on the right-hand side of a formula (32) shows the overflow condition over the decoder buffer 302 \*\*\*\*\* and here.

[Equation 33]

$$R_j \leq N_{max} - N_{j-1} + \bar{R} \quad (33)$$

[Equation 34]

$$R_j \leq \left[ E_{j-L-1} + B_{\max}^d - B_{j-L-1}^{d(e)} - (1 - J_{dec}) R_{j-i} \right] / J_{dec} \quad (34)$$

[Equation 35]

$$N_j = N_{j-1} + R_j - \bar{R} \quad (35)$$

[Equation 36]

$$B_{j-L}^{d(e)} = B_{j-L-1}^{d(e)} + R_j - E_{j-L} \quad (36)$$

By assigning the above-mentioned j value to a formula (36), a value is acquired whenever [ temporary occupancy / of step 502 ].

[0062] We need to do [ as opposed to / when the most /  $j < i+L-1$  ] evaluation only to  $j=i+L-1$ , when [ that ] pre- evaluation can be simplified and reused. However, as shown below,  $B_d(e)_{i-L+1}$  can change, therefore may have reevaluation of all evaluations desirably. Generally, it is good selection that there are more few values of  $R_j \geq 0$  equal to the upper limit of either a formula (32), a formula (33) and a formula (34). However, if the frame which has very many numbers of bits tends to happen, it is known that a smaller value is desirable. A value smaller

than that of  $R_j$  is desirable again, if a channel is the videodisk on which only a few uses a bit as much as possible. If the smaller value of  $R_j$  is chosen, this  $R_j$  will be chosen so that more greatly than the minimum of a formula (32), in order to prevent a decoder buffer underflow.

[0063] To  $i \leq L$ , the frame is not decoded by the video decoder 304 yet, and the decoder buffer 302 is hardly occupied. Generally, in order to avoid the underflow of the decoder buffer 302, the sum of  $R_1, \dots, R_L$  should be chosen so that the expected bit rate of several frames of the beginning of the frame of the bit stream CODEIM of an initial fraction may be exceeded.

[0064] Step 504 evaluates the upper limit (UB) of  $R_{i+L}$  by using the leakage bucket constraint specified by the formula (33).

[Equation 37]

$$R_{i+L} \leq UB \quad R_{i+L} = N_{max} - N_{i+L-1} + \bar{R} \quad (37)$$

At step 505, the upper limit of  $E_i$  is decided by the encoder rate control 209. This upper limit is decided by attaining to overflow of the encoder buffer 203 from inequality (18), and using constraint of underflow \*\* of the decoder buffer 302 from inequality (19).

[Equation 38]

$$E_i \leq B_{\max}^e + R_i - B_{i-1}^e \quad (38)$$

[Equation 39]

$$E_i \leq UBR_{i+L} + B_{i-1}^{d(e)} - J_{\max} UBR_{i+L} \quad (39)$$

The minimum value of these two upper limits of  $E_i$  is supplied by the encoder rate control 209 as an output as a signal RANGE to the video encoder 202.

[0065] At step 506, coding of Frame  $i$  is begun with the video encoder 202 which receives rule of a transfer of the frame from the encoder buffer 203 in the bound and bit rate  $R_i$  of  $E_i$ . At step 507, the transfer which lets coding of the frame  $i$  by the video encoder 202 and the Variable Bit Rate channel 101 of the frame supplied from the encoder buffer 203 pass is completed. The video encoder 202 is here and supplies the real value of  $E_i$  as an input to the channel bit rate control 208 and the encoder rate control 209. The real value of  $E_i$  is used for evaluation of inequality (17) at the conditional-branch point 508, and, thereby, decides whether the underflow of the encoder buffer 203 happened between coding of Frame  $i$ . Control will be passed to step 510 supposing the inspection result in step 508 is NO. Supposing the inspection result in step 508 is YES, termination

of the transfer is carried out early, control is passed to step 509, and the channel rate control 208 and a channel interface 80 are [Equation 40].

$$R_i = B_{i-1}^f + E_i \quad (40)$$

\*\* -- carrying out --  $R_i$  -- being new -- rather than -- a low value -- calculating .

Then, control is passed to step 510. In order to calculate the real value of  $B_{ei}$ , nickel, and  $Bd(e)$   $i-L$  using a formula (8), (35), and (36), respectively, the real value of  $E_i$  and  $R_i$  is used for step 510. At step 511, the increment of the  $i$  is carried out with preparation of coding of the following frame. Supposing step 509 is not performed, all the values before evaluated by step 502 are still effective, and only evaluation of the latest time interval merely needs to be evaluated to the following frame. Supposing step 509 is performed,  $Bd(e)$   $i-L$  changes and all the values before evaluated by step 502 must be reevaluated.

[0066] A simplified block diagram shows the detail of the display-control unit used with either the video presentation controller 303 or the audio control indicator 306 to drawing 6. It is important that only a PTS value is merely used to the audio control indicator 306. A packet header is removed and, specifically, the packet resolver and the DTS (PTS) extractor 601 which detect existence of

the both sides of DTS and PTS or one side are shown. (Since PTS is included, the first received packet is required.) If PTS/DTS is detected (DTS=PTS will be set if it is not DTS), a switch 602 will be set to A location by the signal from a packet resolver and the DTS (PTS) extractor 601, and a DTS value will be supplied to A input of a switch 602. Next, the supplied DTS value is supplied to the DTS (PTS) register 603 as an output by the switch 602 as DTS1.

[0067] After removing a packet header, a packet resolver and the DTS (PTS) extractor 601 pass the remaining packet data by the demand to the presentation unit (PU) detector 607. To the video presentation controller 303, PU is an image expression (namely, picture). To the audio control indicator 306, PU is an audio access unit (AAU). all the first PU(s) (they are not others) -- beginning -- coming out -- the presentation unit detector 607 outputs the clock pulse which carries out the clocking of the clock DTS (PTS) register 603. Thus, the input DTS 1 is read, it memorizes, and a subtractor 604 and an adder 606 are supplied as an output which is DTS2. Next, it waits for the presentation unit detector 607. A Dec\_TC value is supplied to a subtractor 604 and compared with the memorized DTS binary. Difference Dec\_TC-DTS2 produced as a result is supplied to a zero detector 605. If Dec\_TC increases to a value equal to DTS2, the output to the

presentation unit detector 607 and Set B will be supplied for the following PU enabling signal as an input to a switch 602. This permits that the next PU is passed to a decoder 304. In reception of this following PU enabling signal, the presentation unit detector 304 moves the data of the next PU to the video decoder 304 to the video presentation controller 303. The following PU enabling signal makes the data of the next PU similarly transmit to the presentation unit detector 607 to the audio decoder 307 to the audio control indicator 306. The following PU enabling signal makes a switch 602 set to B location again, as stated. If a packet resolver and the DTS (PTS) extractor 601 detect PTS/DTS between the outputs of PU, a switch 602 will be set to A location like a front. However, if PTS/DTS is not detected, a switch 602 stops at B location. In the end of PU data transfer by the presentation unit detector 607, this outputs a clock pulse to the clocked into of the DTS (PTS) register 603 like a front. If a new DTS value comes, it will be read like a front from A input of a switch 602, will be supplied to a register 603 as DTS1 value, and will appear as a DTS output value from a register 603. However, if a new DTS value does not come, a switch 602 will still stop at B location. In this situation, a presumed DTS value is supplied from an adder 606, is supplied as DTS1 new value to a register 603, and

appears as that output as DTS2. This presumed DTS value is acquired by applying a temporary fixed picture duration value to the front DTS binary per Dec\_TC with an adder 606. For example, a constant is 3003, if a temporary picture frequency is 29.97Hz and a Dec\_TC frequency is 90Hz. Next, the presentation unit detector 607 receives the following PU enabling signal (pulse) from a zero detector 605 waiting and here until Dec\_TC increases to the new value of this DTS2. Thus, a data output cycle is repeated.

[0068]

[Effect of the Invention] As stated above, the video signal coding equipment of this invention The encoder / decoder buffer overflow, and the underflow which encounter when actually using the channel of a strange good bit rate adjustable or effectually, in order to communicate the audio signal which corresponds [ which corresponds and coding-video-imagines ] The expression of cel delay change for which it opted in the decoder which separated, i.e., a jitter, can be answered, it can conquer by adjusting the parameter of a video encoder, and highly efficient video encoder / decoder which forces it a real-time cutting tool delivery schedule to a real decoder can be offered.

---

## DESCRIPTION OF DRAWINGS

---

[Brief Description of the Drawings]

[Drawing 1] The simplified block diagram of the multiplexer unit which took in the Variable Bit Rate channel and demultiplexer unit of this invention is shown.

[Drawing 2] The detailed simplified block diagram of the multiplexer unit of

drawing 1 is shown.

[Drawing 3] The detailed simplified block diagram of the demultiplexer unit of drawing 1 is shown.

[Drawing 4] It is the flow chart showing actuation of the count jitter unit of drawing 3.

[Drawing 5] It is the flow chart showing the video encoder buffer control by which a Variable Bit Rate channel is restrained effectually.

[Drawing 6] The detailed simplified block diagram of the unit in which drawing 3 carries out video presentation is shown.

[Description of Notations]

100 Multiplexer Unit

101 Variable Bit Rate Channel

102 Demultiplexer Unit

200 Multiplexer Unit

201 DTS Unit

202 Video Encoder

203 Video Presentation Control Unit

204 Insertion PTS Unit

205 Audio Encoder  
206 Audio Data Buffer  
208 Channel Rate Control  
209 Encoder Rate Control  
210 System Encoder, Enc\_CR Aedeagus, and Jdec Extractor  
212 System Time Clock Generation Machine  
214 Oscillator  
215 Output Counter  
300 Demultiplexer Unit  
301 System Decoder, Enc\_CR Extractor, and Jdec Aedeagus  
302 Video-Data Buffer  
303 Video Presentation Controller  
304 Video Decoder  
305 Audio Data Buffer  
306 Audio Control Indicator  
307 Audio Decoder  
308 System Time Clock Generation Machine  
310 Phase Comparator

311 Filter

312 VCO

314 Output Counter

317 Count Jitter Unit

401 Initiation

402 Set Starting Value of Jitter Jmax.

403 Obtain Enc\_CR-Dec\_TC.

404 Calculate Absolute Value  $J = |Enc\_CR-Dec\_TC|$ .

405  $J < Jmax - \delta$  ?

406  $Jmax < J$  ?

407 Decrease Jmax by  $\Delta/2$ .

408 Set to  $Jmax = J$ .

409 Output Value of  $Jmax = Jdec$ .

500 Initiation

504 Evaluate Upper Limit of  $R_{i+1}$ .

505 Evaluate Upper Limit of  $E_i$ .

506 Start Transfer for Coding of Frame I at Initiation and Rate  $R_i$ .

507 Acquire Actual Value of Termination and  $R_i$  for Coding and Transfer.

508 Is Underflow of the Encoder Buffer Carried Out?

509 Re-calculate Actual Value of Ri.

511 Increment I. To the Following Frame

601 Packet Resolver and DTS (PTS) Extractor

602 Switch

603 DTS (PTS) Register

604 Subtractor

605 Zero Detector

606 Adder

607 Presentation Unit (PU) Detector